The Dissertation Committee for Wei Wu
certifies that this is the approved version of the following dissertation:

Networked Control and Efficient Transmission in Sensor Networks

Committee:

Ari Arapostathis, Supervisor

Gustavo De Veciana

Sanjay Shakkottai

Vijay Garg

David Morton

Networked Control and Efficient Transmission in Sensor Networks

by

Wei Wu, M.S.; B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2007

To my parents and Ning

Acknowledgments

I would like to thank my advisor Professor Ari Arapostathis for his encouragement, his advisement and his generous research spirit throughout my PhD study. There have been countless times that we had long discussion till midnight, during which countless coffee have been drunken at JP coffee. I benefit a lot not only from his broad knowledge and deep insight in engineering and mathematics, but also his research style to approach the solutions.

I would like to thank Professor Sriram Vishwanath, not only for his instruction in the area of information theory, but also personal suggestions on my career development. We have spent a lot of time in discussing research problems in feedback capacity and cognitive radio channels, as well as improving presentation skills and publicizing the research results. Working with him has been an invaluable and pleasant experience for me.

I would also like to thank Professor Sanjay Shakkottai, Professor Gustavo de Veciana, Professor Vijay Garg, and Professor David Morton for serving on my committee, and for their inspiring suggestions toward my PhD work. I thank Professor Constantine Caramanis for his suggestions on my presentation.

I want to thank my labmates Keethi Nagaraj, Johnson Carroll, Tom Rosenwinkel, Dafeng Wang, Chun-hung Liu, Youngsang Kim, who gave me a lot of help, and colleagues at WNCG Sundar Subramanian, Chulhan Lee, Caleb Lo, Brian Smith, Sriram Sridharan, who were always willing for discussion.

I also want to thank all students at WNCG who make a great research atmosphere and a pleasant environment.

Finally, I would like to express my sincere thanks to my wife Ning and my parents for their love, encouragement, and support.

WEI WU

The University of Texas at Austin
August 2007

Networked Control and Efficient Transmission in Sensor Networks

Publication No. ____________________

Wei Wu, Ph.D.
The University of Texas at Austin, 2007

Supervisor: Ari Arapostathis

Enabling "intelligent environments" that are effortlessly automated is a key promise of sensor networks of the future. These networks have a wide range of domains in which they can be effectively deployed, including health-care, emergency response, manufacturing and surveillance. Unlike the majority of existing (and perhaps better-understood) network configurations, wireless-implemented sensor networks suffer from extremely stringent constraints in terms of size, battery power and computing ability, and possess distinctive features in terms of scalability and end-goal of deployment. Thus, it is imperative that we determine solutions that are tailored to the constraints and goals of these systems, by bringing together ideas in the domains of control, computing and communications to a common analytical platform.

In this dissertation, we build a theoretical framework that uses system theory, stochastic control, queuing theory and information theory to determine the following:

1. A characterization of the stability and optimal control policies with sensor querying (i.e. which set of sensors must be queried and when) using system theory and stochastic control;
2. A delay-optimal energy efficient transmission scheme for these networks (i.e. with what power level must they communicate) using heavy traffic limits and stochastic control; and
3. a cooperative transmission strategy for maximizing capacity of these networks (i.e. how they should encode their data to send the most through) using network information theory.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

In recent years, sensor networks have emerged as a common platform for control, computing and communications to support a wide spectrum of engineering applications, including monitoring, surveillance, fault detection, process control, distributed computing and mobile communications.

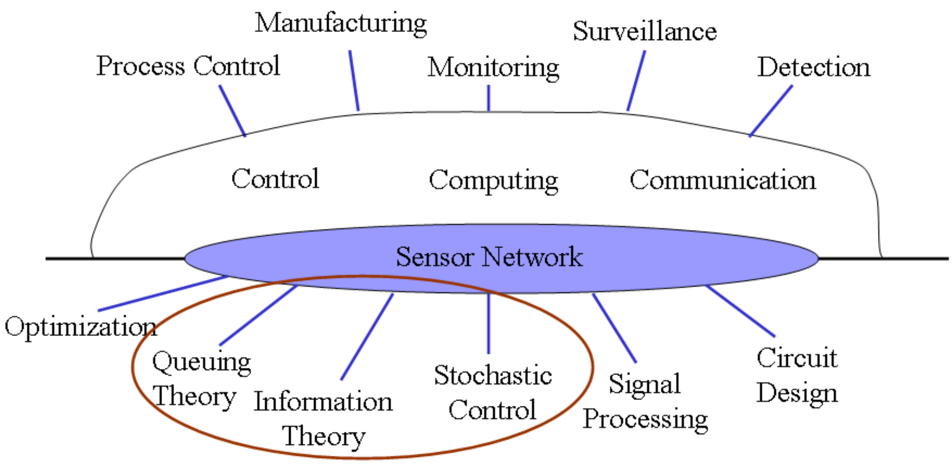


Figure 1.1: Sensor networks: a common platform for control, comouting and communications

On the other hand, the convergence of control, computing and communications into one common device and/or platform has resulted in a plethora of research challenges. As an example of this convergence, consider an automated intelligent factory equipped with hundreds of sensors, with each sensor sensing multiple processes (temperature, motion, chemical) simultaneously, thus monitoring the overall environment. Another example is the on-the-fly deployment of sensors/mobile devices in a military scenario. Key research issues that govern this networked system include: system detectability and stability, data collection, time-delay, transmission efficiency, et al. Moreover, the cross-disciplinary nature of research challenges in this context calls for collaborative research in different areas, including circuit design (for low-power devices), signal processing, optimization, queueing theory, information theory and stochastic control theory.

In this dissertation, we will take views from both system theory and information theory and deal with challenges arising in a networked control system, a control system enabled with a network of distributed sensors. For example, Figure 1.2 shows such a system with a centralized controller (or processing center): distributed sensors acquire local information from their environment, make some preprocessing, e.g., quantization, compressing and encoding, and send the data to the processing center through the wireless sensor network; the controller (or processing center) collects data and makes decisions according to the objectives of the system design; and finally the decisions (or control commands) are sent to distributed actuators again through the sensor network, and actuators exercise the control afterward.

In this feedback loop, the only difference from a classical control system is that the observation and the control are required to be transmitted through a network other than perfectly known. However, the nature of sensor networks results in research challenges at different levels: at the system level, the large-scale deployment of sensors call for a distributed design of the system, which is able to scale with the size of the network; at the network level, the system is built by a network interconnected by wireless links that suffers interference and network congestion and is subject to stringent delay constraint due to control applications; at the individual node level, the physical environment and the system scale limit sensors to be cost-effective low power devices that can survive with their limited energy and computing capability.

As a result, we need to take a cross-layer look at the network and tackle these challenges by bringing together ideas from systems and control, communications and networking. In this dissertation, we mainly consider three issues arising at different

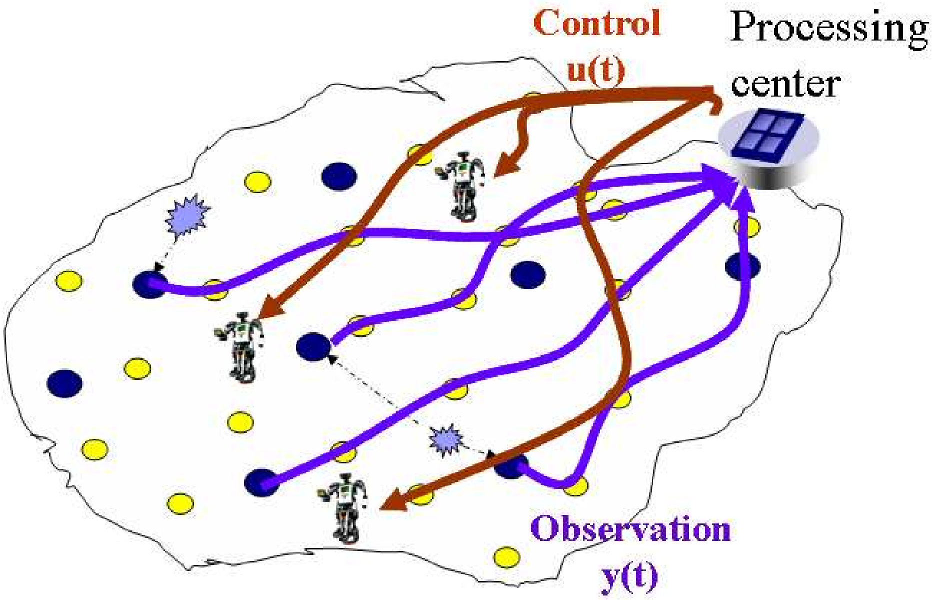


Figure 1.2: A networked control system

levels: (i) more information at the processing center leads to better decision making but results in heavy burden on the network transmission, thus we study optimal sensor querying in a networked control system to balance the information required at the processing center and the system performance; (ii) since opportunistic transmission, i.e., transmitting only when channel quality is high, is definitely the best way to achieve high energy efficiency but might induce large delay, we study the power-delay trade-off and establish the optimality of a simple threshold policy in large delay asymptotics; (iii) data correlation among transmitters can potentially boost the transmission rate even in an interference-limited channel, and here we model it as an interference channel with one transmitter knowing another's message and establish its capacity region using the tools of network information theory.

1.2 Contents and organization

1.2.1 Chapter 2: preliminaries - ergodic control of Markov decision processes

This chapter provides some preliminary background in the area of ergodic control of Markov decision processes (MDP) that will be further developed upon in later chapters. Markov decision processes are a class of sequential decision making problems in which the system dynamics evolve in a Markovian manner. Ergodic control is the class of MDP problems that minimize the long time average cost, while the optimal cost is independent of the initial condition. In the existing literature, the varnishing discount method has been widely adopted to characterize the optimal control for the ergodic cost, and the convex analytic method appears to be a powerful tool to show the existence of ergodic control in both discrete-time MDP models and continuous-time MDP models. We summarize the varnishing discount method and the convex analytic method and specialize to partially observable MDP problems in discrete-time models, and ergodic control with constraints in continuous-time models.

1.2.2 Chapter 3-4: optimal sensor querying for networked control systems

In these two chapters, we consider a system-level problem: given the vast amount of information available, how to design sensor querying policies such that the control system achieves certain objectives, e.g., stability or minimizing some system performance metrics. In a large-scale networked control system, controllers require timely information from sensors for decision-making. It is impossible to collect the data from all the sensors and send them to the controllers all the time, since it will induce tremendous information traffic, that consumes already limited resources and causes large delay. An alternative way is to query a small group of sensors at a time. However, this is not only at the expense of performance, but might also result in the loss of stability of the system. Moreover, there might be a cost associated with query, representing the energy consumption for data collection or the traffic induced in the network. The controllers have to make a tradeoff between their control objectives and the querying cost. It is crucial to understand how the distributedness of the sensors affects the stability and what the optimal querying law given a network configuration and the querying cost structure. This problem is formulated under the general framework of POMDP in Chapter 3 and the existence of ergodic control is established under a mild assumption for a hierarchical information structure.

In Chapter 4, we study the optimal sensor querying in the context of the linear quadratic Gaussian (LQG) control problem. In this model, the controller has a set of querying options to choose from and can only select one of these options at each time. We establish the necessary and sufficient condition for stability and characterize the infinite-horizon optimal control for the discounted and average costs. One of the key results is that *distributing the sensors does not hurt stability*, in other words, as long as the system can be stabilized with all the sensor data available at the controller, there exists a querying rule that can stabilize the system as well. We also obtain sharp conditions for the existence of ergodic control for the average cost problem, and show that the separation principle partially holds. This provides the justification for the use of sensor querying in a large-scale sensor network and allows for the separate design of estimation and control.

1.2.3 Chapter 5: optimal transmission in a time-varying channel - heavy traffic analysis

Once we have established a decision rule for sensor querying, i.e., from which sensors the processing center should collect information, the data needs to be delivered in a timely manner. End-to-end delay is of paramount importance for networked control systems, otherwise they might fail catastrophically. On the other hand, transmitters have to combat with time-varying wireless fading and limited battery. Thus it is critical to characterize the optimal policy for power allocation under delay constraints. Previous work in the literature models the problem as a discrete-time constrained Markov decision process and solves the dynamic programming equation numerically, which requires tremendous computation effort and lacks insight into the nature of the problem. In Chapter 5, by proper time scaling of channel fading and the queueing process, we have applied heavy traffic analysis to obtain the diffusion approximation of the queueing process, which can finally lead to an ergodic control problem for a diffusion process and yield close-form solutions. Our result shows that a channel-state dependent threshold policy is optimal to achieve power-delay tradeoff. Besides, on the theory front, we have also made substantial progress in ergodic control of diffusion processes with constraints, which might have broader impacts on solving other engineering problems.

1.2.4 Chapter 6: cognitive transmission- information theoretic view on node cooperation in interference channels

Power allocation can improve power efficiency under delay constraints only for a transmitter-receiver pair. In a densely-deployed sensor network, interference among neighboring nodes can severely hurt transmission efficiency. Thus the information transmission can only be improved if it can be done collaboratively rather than independently.

In Chapter 6, we take an information theoretic view on the cognitive transmission and node cooperation in a interference environment. On one hand, "smart" transmitters can learn from the environments and recognize messages transmitted by interfering neighbor sensors, thus inducing certain level of node cooperation. On the other hand, nodes close by might share common information due to the correlation in the sensor field so as to cooperate rather than interfere with each other. We have modeled both situations as *an interference channel with degraded message sets* and established the capacity region of this class of channels, including both discrete memoryless channels and Gaussian additive interference channels: for discrete memoryless channels, single letter characterizations of the inner bound and outer bound are obtained and they match together under a couple of weak interference assumptions; for Gaussian cases, both single letter characterization and Gaussian input optimality are obtained in the weak interference scenario that yields a close-form solution for the capacity region. These results from the information theoretic perspective shed some light on the performance limit of cognitive transmission in an interference channel, which has been considered as the future of wireless technologies and has been proposed for a more efficient use of the radio spectrum.

Chapter 2

Preliminaries: Ergodic Control of Markov Decision Processes

2.1 Introduction

A Markov decision process (MDP), also called a controlled Markov process, is a sequential decision problem in which the state evolves in a Markovian manner given the history of controls and states, and the objective is to minimize the sum of a given running cost in a finite or infinite horizon. In this dissertation, we focus on a class of MDP problems, called *ergodic control* problems, that minimizes the limiting time-averaged cost of the system. Ergodic control problems are particularly important in applications to communication systems, where a "steady state" operation is expected over intervals that are long compared to the time constant of the system.

Mathematically, the ergodic cost criterion stands out as being much more difficult to analyze than the others (e.g., discounted cost criterion and finite-horizon cost criterion). Even though the discrete-time MDP problem with finite states and actions is well understood, its partially observed counterpart still suffers from analytical and computational difficulties [81] [33] [12]. The continuous-time MDP problem, e.g., ergodic control of diffusions, is even more technically involved than the discrete-time MDP problem. A variety of approaches have been developed to handle different situations. For example, the approaches for discrete-time MDP are summarized in [3] and those for continuous-time MDP are discussed in [14] [17].

In this chapter, we briefly summarize some of new development on the ergodic control of POMDP [50] and controlled diffusions [4], which are applied to the analysis in Chapter 3, Chapter 4 for discrete-time and Chapter 5 for controlled diffusions.

2.2 Discrete-time MDP

We consider a controlled Markov chain $(\boldsymbol{X}, \boldsymbol{U}, P, \mu, g)$ where $\boldsymbol{X}$ is the state space, $\boldsymbol{U}$ is the action space, μ is the distribution of the initial state X_0, and $g(x, u)$ is the one-stage cost that is incurred when the system is at state x and action u is applied. The *history spaces* are defined as

$$\boldsymbol{H}_0 := \boldsymbol{X}\,, \qquad \boldsymbol{H}_t := \boldsymbol{H}_{t-1} \times (\boldsymbol{X} \times \boldsymbol{U})\,, \quad t \in \mathbb{N}_0\,.$$

An *admissible control strategy*, or *policy*, is a sequence $v = \{v_t\}_{t \in T}$ of Borel measurable stochastic kernels on $\boldsymbol{U}$ given $\boldsymbol{H}_t$, satisfying the constraint

$$v_t\big(U(x_t) \bigm| h_t\big) = 1\,, \qquad x_t \in \boldsymbol{X}\,, \quad h_t \in \boldsymbol{H}_t\,.$$

The set of all admissible policies is denoted by $\mathit{\Pi}$. Note that the admissible policies are non-anticipative policies, i.e., do not depend on the future but only the history. $\mathbb{P}^v_\mu$ is the probability measure under policy v with the initial condition $X_0 = x$. The expectation operator with respect to $\mathbb{P}^v_\mu$ is denoted by $\mathbb{E}^v_\mu$. Furthermore, if μ is a Dirac measure at $x \in \boldsymbol{X}$, we simply write $\mathbb{P}^v_x$ and $\mathbb{E}^v_x$. The following criteria are frequently used.

- Finite-horizon cost. The total cost incurred by the policy $v \in \mathit{\Pi}$ over the entire planning horizon is given by

$$\begin{aligned} J_N(\mu, v) &\triangleq \mathbb{E}^v_\mu \left[\sum_{t=0}^{N-1} g(X_t, U_t) + G(X_N) \right] \\ J^*_N(\mu) &\triangleq \inf_{v \in \mathit{\Pi}} J_N(\mu, v)\,, \end{aligned}$$

  where G is the terminal cost.

- Discounted cost. Let $0 < \beta < 1$, the *discount factor,* and $v \in \mathit{\Pi}$ be given. The total discounted cost incurred by v over the infinite planning horizon is given by

$$J_\beta(\mu, v) \triangleq E_\mu^v \left[\sum_{t=0}^{\infty} \beta^t g(X_t, U_t) \right]$$
$$J_\beta^*(\mu) \triangleq \inf_{v \in \Pi} J_\beta(\mu, v)\,.$$

- Average cost. The expected long-run average cost incurred by $v \in \Pi$ is given by

$$J(\mu, v) \triangleq \limsup_{N \to \infty} \mathbb{E}_\mu^v \left[\frac{1}{N} \sum_{t=0}^{N-1} g(X_t, U_t) \right]$$
$$J^*(\mu) \triangleq \inf_{v \in \Pi} J(\mu, v)\,.$$

- Sample path average cost. This is a path-wise version of the AC, and, for $X_0 = x$, it is given by

$$J_S(x, v) \triangleq \limsup_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N-1} g(X_t, U_t)$$
$$J_S^*(x) \triangleq \inf_{v \in \Pi} J_S(x, v)\,.$$

where $\{X_t\}$ and $\{U_t\}$ are the state and control process induced by $v \in \Pi$.

Define the *dynamic programming map* T by

$$T(J) \triangleq \inf_{u \in \boldsymbol{U}} \left\{ g(x, a) + \int_{\boldsymbol{X}} J(x') P(dx' \mid x, u) \right\} \qquad \forall\, x \in \boldsymbol{X} \tag{2.1}$$

According to dynamic programming principle, the optimal solution of the finite-horizon cost can be characterized by

$$J_t^* = T(J_{t+1}^*)\,.$$

Define the *discounted dynamic programming map* T_β by

$$T_\beta(J) \triangleq \inf_{u \in \boldsymbol{U}} \left\{ g(x, a) + \beta \int_{\boldsymbol{X}} J(x') P(dx' \mid x, u) \right\} \qquad \forall\, x \in \boldsymbol{X}$$

The characterization for the discounted cost can be obtained through the finite-horizon cost under additional assumptions on stability: for example, if $\mathbb{E}_\mu^v|g(X_t, U_t)|$ is bounded, the optimal solution of the discounted cost can be proved to be the limit of a series of optimal solutions of the finite-horizon cost, and is determined by the *discounted cost optimality equation* (DCOE)

$$J_\beta^* = T_\beta(J_\beta^*)\,,$$

For a bounded running cost g, the existence of a unique solution to the DCOE can be established via the contraction mapping theorem.

Although the average cost criterion can be treated as the limiting case of the finite horizon problem, it is often better to treat it as the limiting case of the discounted cost criterion as $\beta \to 1$, which is called the *vanishing discount* approach. Under certain conditions, the optimal average cost is characterized by the *average cost optimality equation*,

$$J^* + h = T(h)\,,$$

where J^* is a scalar denoting the optimal average cost. Since the optimal cost is independent of the initial condition $X_0 = x$, it is also called *ergodic control* and the average cost is equal to the sample-path average cost.

One approach to prove the existence of a solution for the ACOE is to construct the differential discounted cost value function as

$$\hat{h}_\beta(x) \triangleq J_\beta^*(x) - \inf_{x\in\boldsymbol{X}} J_\beta^*(x)\,, \tag{2.2}$$

and show that

$$\begin{aligned} J^* &= \lim_{\beta\to 1}(1-\beta)\inf_{x\in\boldsymbol{X}} J_\beta^*(x) \\ h &= \lim_{\beta\to 1}(1-\beta)\hat{h}_\beta(x)\,, \end{aligned}$$

for some sequence of discounted cost value function $\{J_\beta^*\}$.

For a POMDP, the decision maker (or controller) cannot fully access the state of the Markov chain, but is implied with an observation process $\{Y_t\} \in \boldsymbol{Y}$, e.g. $Y_t = f(X_t)$. In general, the dynamics of the process are governed by a transition kernel on $\boldsymbol{X} \times \boldsymbol{Y}$ given $\boldsymbol{X} \times \boldsymbol{U}$, which may be interpreted as

$$Q_{ij}(y,u) \triangleq \text{Prob}(X_{t+1} = j, Y_{t+1} = y \mid X_t = i, U_t = u)\,.$$

It is well known that for a POMDP model, one can derive a completely observed (CO) model which is equivalent to the original model in the sense that for every control policy in the POMDP model there corresponds a policy in the CO model that results in the same cost, and vise versa.

Consider a POMDP with finite state space $\boldsymbol{X}$, compact action space $\boldsymbol{U}$ and history spaces

$$\boldsymbol{H}_0 := \boldsymbol{X}\,, \qquad \boldsymbol{H}_t := \boldsymbol{H}_{t-1} \times (\boldsymbol{Y} \times \boldsymbol{U})\,, \quad t \in \mathbb{N}_0\,.$$

The state of the equivalent CO model, Ψ_t, is the state distribution of the POMDP model given the history, namely,

$$\Psi_t \triangleq \mathbb{E}[X_t \mid h_t]\,.$$

The transition of the system dynamics is determined by the Baysian filter,

$$\psi_{t+1} = \mathcal{T}(\psi_t, y_t, u_t)\,,$$

where $\mathcal{T}$ is defined by

$$\begin{aligned} V(\psi, y, u) &= \psi Q(y,u)\mathbf{1} \\ \mathcal{T}(\psi, y, u) &= \begin{cases} \frac{\psi Q(y,u)}{V(\psi,y,u)} & \text{if } \; V(\psi,y,u) > 0 \\ \frac{1}{n}\mathbf{1} & \text{otherwise}\,. \end{cases} \end{aligned}$$

After this transformation, a POMDP can be treated a CO MDP problem with new state space $[0,1]^n \subset \mathbb{R}^n$ and bounded cost. For this class of problems, there always exists a unique solution for the DCOE, thus the discounted cost problem for a POMDP is well understood. However, for the average cost, there are numerous examples for which the associated ACOE has no solution. Indeed, checking the unichain condition (existence of solution for the ACOE with constant J^*) for a general MDP has recently been shown to be NP-hard.

There have been many efforts to derive sufficient conditions for existence of solutions to the ACOE for POMDPs. Most of the well known sufficient conditions, including Ross's *renewability* condition [84], Platzman's *reachability-detectability* condition [80], and Stettner's *positivity* condition [86], are not general enough and fail in even some simple problems that are known to possess stationary optimal policies. Recently in [50], a so-called *interior accessibility* condition, which is more general than the conditions mentioned above [80, 84, 86], has been proven to guarantee existence.

Assumption 2.2.1 (Interior accessibility)**.** *Define*

$$\mathbf{\Psi}_\epsilon = \{\psi \in \mathbf{\Psi} : \psi(i) \geq \epsilon, \forall i \in \boldsymbol{X}\}\,.$$

There exist constants $\epsilon > 0$, $k_0 \in \mathbb{N}$ and $\beta_0 < 1$ such that if $\psi_*^\beta = \arg\min_{\psi\in\mathbf{\Psi}} J_\beta^*(\psi)$, then for each $\beta \in [\beta_0, 1)$, we have

$$\max_{1\leq k\leq k_0} \mathbb{P}^{v_\beta}_{\psi_*^\beta}(\varPsi_k \in \mathbf{\Psi}_\epsilon)\,,$$

where v_β denotes the optimal policy for β-discounted cost and J_β^* is its value function.

The verification of this condition is simplified due to the fact that J_β^* is concave and its minimum ψ_*^β is attained on the extreme points of $\gneqq$. Some key ideas behind the proof are invoked in the proof of Theorem 4.5.5, the main theorem in Chapter 4. Assumption 2.2.1 is stated in term of the optimal discounted policy v_β, which might not be directly available. However, it can be replaced by a stronger condition below, which only depends on the transition kernels.

We adopt the following notation. For $k \in \mathbb{N}$, let

$$\begin{aligned} u^k &\triangleq (u_0, \dots, u_{k-1}) \in \boldsymbol{U}^k \\ y^k &\triangleq (y_1, \dots, y_k) \\ Q(y^k, u^k) &\triangleq Q(y_k, u_{k-1}) \cdots Q(y_2, u_1) Q(y_1, u_0)\,. \end{aligned}$$

Assumption 2.2.2. *There exists $k_0 \in \mathbb{N}$ such that, for each $i \in \boldsymbol{X}$,*

$$\max_{1\leq k\leq k_0} \min_{u^k\in\boldsymbol{U}^k} \left\{ \max_{y^k\in\boldsymbol{Y}^k} \min_{j\in\boldsymbol{X}} Q_{ij}(y^k, u^k) \right\} > 0\,. \tag{2.3}$$

Perhaps a more transparent way of stating Assumption 2.2.2 is that for each $i \in \boldsymbol{X}$, and for each sequence $u^{k_0} = (u_0, \ldots, u_{k_0-1})$, there exists some $k \leq k_0$ and a sequence y^k, such that $Q_{ij}(y^k, u^k) > 0$, for all $j \in \boldsymbol{X}$. This result is applied to prove Theorem 3.4.1 later in Chapter 3.

2.3 Continuous-time MDP

Among many continuous-time MDP models, controlled diffusion processes are perhaps the most important and are widely applied to communication networks, manufacturing systems, financial engineering and many other fields. Diffusion processes are a class of random processes on $\mathbb{R}^n$ driven by Brownian motions or *Wiener processes*. The prototypical controlled diffusion process $\{X_t, t \geq 0\}$ can be described by a stochastic differential equation (SDE)

$$dX_t = b(X_t, U_t)dt + \sigma(X_t)dW_t\,. \tag{2.4}$$

Here,

1. U_t takes value in a compact metric space $\boldsymbol{U}$ and has measurable sample paths. In addition, it is *non-anticipative.*

2. $b(x, u)$ is Lipschitz in x uniformly respect to u, and $\sigma(x)$ is Lipschitz.

3. W_t is a d-dimensional standard Wiener process independent of X_0.

The class of $U_t \in \boldsymbol{U}$ enunciated above is the most general class of controls, which only requires to be non-anticipative controls. Let $\{\mathcal{F}_t^X\}$ denote the natural filtration of X_t. U_t is a feedback control if it is adapted to $\{\mathcal{F}_t^X\}$, i.e., U_t at each time t is a function of the observed trajectory $X([0,t])$. It is a Markov control if in addition $U(t) = v(t, X_t)$, for a measurable function v. Finally, it is a stationary Markov control if $U_t = v(X_t)$ for a measurable function v.

We shall need a relaxation of the notion of control process above to that of a *relaxed control* process. The idea of relaxed controls is a continuous-time version of randomized policies. The new control space becomes $\mathcal{P}(\boldsymbol{U})$, a space of probability measures on $\boldsymbol{U}$. Moreover,

$$\bar{b}(x, v) = \int_{\boldsymbol{U}} b(x, u)v(du)\,,$$

which inherits the Lipschitz conditions from b, and $\bar{\sigma}(x, v)$ is defined as the nonnegative definite square root of

$$\int_{\boldsymbol{U}} \sigma(x, u)^T \sigma(x, u) v(du) \, .$$

The original notion of controls in the space $\boldsymbol{U}$ is then a Dirac measure in $\mathcal{P}(\boldsymbol{U})$, which is called *precise controls*. The benefit of the set of relaxed controls is that $\mathcal{P}(\boldsymbol{U})$ is the compactification and convexization of the set of precise controls, $\boldsymbol{U}$, and every precise control corresponds to an extreme point of the compact convex set.

Like their discrete-time counterparts, the objective of continuous-time optimal control problems is to minimize some standard cost functionals:

- Finite-horizon cost: For $T > 0$, minimize

$$\mathbb{E}\Big[\int_0^T c(X_t, U_t)dt + C(X_T)\Big] \, ,$$

  where c is the running cost and C is the terminal cost.

- Infinite horizon cost: For a discounted factor $\alpha > 0$, minimize

$$\mathbb{E}\Big[\int_0^\infty \exp(-\alpha t) c(X_t, U_t)dt\Big] \, .$$

- Average cost: minimize

$$\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\Big[\int_0^T c(X_t, U_t)dt\Big] \, .$$

For a Markov process X_t, its generator is defined as

$$\mathcal{L}f(x) = \lim_{t \to 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t} \, .$$

For the controlled diffusion (or SDE) defined in (2.4), its generator with control u is

$$\mathcal{L}^u f(x, u) = \frac{1}{2} \sum_{ij} \sigma_i j(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_i b_i(x, u) \frac{\partial f}{\partial x_i}(x) \, .$$

2.3.1 Existence of optimal control

Next we describe the convex analytic approach. Let $v(\cdot)$ be a stationary Markov control such that corresponding X_t is positive recurrent and there has a unique occupation measure as

$$\pi_v(dx, du) = \eta_v(dx)v(x, du)\,,$$

where $\pi \in \mathcal{G}$ is the set of all ergodic occupation measures, $\eta_v \in \mathcal{M}$ is the set of all invariant probability measures, $v(x, du)$ is the randomization of the relaxed control v with some abuse of notation, and η_v is the invariant probability measure corresponding to the stable control v.

It can be shown that η is an invariant measure under control v if and only if

$$\int_{\mathbb{R}^n} \mathcal{L}^v f d\eta_v = 0\,,$$

for any function $f \in \mathcal{C}^2(\mathbb{R}^n)$ (space of functions with continuous derivatives up to the second order) with compact support. A similar result can be directly extended to the ergodic occupation measure: a probability measure π is an ergodic occupation measure if and only if

$$\int_{\mathbb{R}^n} \mathcal{L}^u f d\pi_v(dx, du) = 0\,,$$

where π_v can be decomposed as $\pi_v(dx, du) = \eta_v(dx)v(x, du)$. The following theorem holds.

Theorem 2.3.1. *The set $\mathcal{G}$ of all ergodic occupation measures is closed and convex, and its extreme points corresponds to the precise controls.*

Applying Choquet's theorem, the important outcome due to this structure of $\mathcal{G}$ is that all $\pi \in \mathcal{G}$ is the barycenter of a probability measure supported on $\mathcal{G}_e$, the set of all extreme points of $\mathcal{G}$. Moreover, considering the cost function $\int c d\pi$, this becomes a linear program on a convex set, namely,

$$\begin{aligned} \text{minimize:} \quad & \int_{\mathbb{R}^n \times \boldsymbol{U}} c d\pi \\ \text{over} \quad & \pi \in \mathcal{G}\,. \end{aligned}$$

If the minimum is attained, it is attained at certain extreme point that corresponds to certain stationary stable Markov control.

However, the minimum might not be attained. For example, if $c(x) = \exp(-|x|^2)$, the ergodic cost for all stable Markov controls is positive, while unstable Markov control can have cost 0, thus making the later optimal. In practical applications, we want to rule out such scenarios and focus on problems where stability and optimality are not at odds. For this reason, the existing literature imposes two different sets of assumptions: (a) a condition on the cost function that penalizes unstable behaviors; and (b) a blanket stability assumption on the SDE (2.4):

Assumption 2.3.1 (near-monotone)**.** *The cost function c is* near-monotone*, namely,*

$$\lim_{|x|\to\infty} \inf_{u\in\boldsymbol{U}} c(x,u) > \rho^* \,, \tag{2.5}$$

where

$$\rho^* \triangleq \inf_{\pi\in\mathcal{G}} \int_{\mathbb{R}^n\times\boldsymbol{U}} c d\pi \,.$$

Assumption 2.3.2 (stable)**.** *$\mathcal{G}$ is compact.*

It may seem that (2.5) is difficult to verify because ρ^* is usually unknown. However, there are important cases in which Assumption 2.3.1 automatically holds. One example is that $\inf_{u\boldsymbol{U}} c(x,u)$ grows unbounded when $|x| \to \infty$ and $\rho^* < \infty$. Another example is the case when $c(x,u) = c(x)$ does not depend on u and $c(x) < \liminf_{|y|\to\infty} c(y)$. Both examples can cover applications in queueing systems.

With either Assumption 2.3.1 or Assumption 2.3.2, the existence of ergodic control can be shown.

Theorem 2.3.2. *Under either Assumption 2.3.1 or Assumption 2.3.2, $\int c d\pi$ attains its minimum in $\mathcal{G}$ at some $\pi^* \in \mathcal{G}$; moreover, π^* corresponds to a policy that is stable, stationary, deterministic, and Markovian.*

A similar result has been extended to a class of ergodic control problems with constraints [16, 18]. The objective is to minimize

$$\int_{\mathbb{R}^n\times\boldsymbol{U}} c_0 d\pi \,,$$

over $\pi \in \mathcal{G}$, subject to

$$\int_{\mathbb{R}^n \times \boldsymbol{U}} c_i d\pi \leq \overline{m}_i \,, \quad 1 \leq i \leq l \,.$$

Define ρ^* as the infimum of average cost that can be achieved by stable Markov controls under the constraints. The existence of ergodic control can be proved under the following extended near-monotone assumption.

Assumption 2.3.3.

$$\liminf_{|x|\to\infty} \inf_{u\in\boldsymbol{U}} c_0(x,u) > \rho \qquad \liminf_{|x|\to\infty} \inf_{u\in\boldsymbol{U}} c_i(x,u) > \overline{m}_i \,, \qquad 1 \leq i \leq l \,.$$

However, in the problem we consider in Chapter 5, the above assumption is not satisfied. Thus a new technique applying Lagrangian multipliers is introduced to establish the existence of ergodic control with constraints.

On the other hand, optimality in Theorem 2.3.2 is over all stable stationary controls. Indeed, the solution in Theorem 2.3.2 can be proven optimal among all admissible controls as well, including nonstationary ones.

2.3.2 Characterization of optimal control

In this section, the optimal policy of ergodic cost criterion is characterized in term of the solution of certain Hamilton-Jacobi-Bellman (HJB) equation. For a stable stationary Markov control v, let

$$\rho_v \triangleq \int_{\mathbb{R}^n} c(x,v)\eta_v(dx) \,,$$

where η_v is the unique invariant probability measure corresponding to v, and

$$\rho^* \triangleq \inf_{v\in\mathfrak{U}_{\mathrm{ss}}} \rho_v \,,$$

where $\mathfrak{U}_{\mathrm{ss}}$ denotes the set of stable stationary Markov controls. As discussed in Section 2.3.1, the infimum can be attained in $\mathfrak{U}_{\mathrm{ss}}$ and its minimum is optimal in all admissible controls

Similar the approach taken in the discrete-time MDP, the characterization of optimality is established through vanishing discount method by taking the limit of the HJB equations for the discounted cost criterion, as the discount factor approaches zero.

For a discounted factor α and an admissible control $U \in \mathfrak{U}$, the α-discounted cost is defined as

$$J_\alpha^U \triangleq \mathbb{E}_x^U \Big[\int_0^\infty e^{-\alpha t} c(X_t, U_t) dt \Big] ,$$

and the optimal cost is defined as

$$V_\alpha(x) \triangleq \inf_{U \in \mathfrak{U}} J_\alpha^U(x) .$$

It is known that for a bounded cost $c(x,u)$, which is Lipschitz in x uniformly to u, $V_\alpha(x)$ is the unique solution of

$$\inf_{u \in \boldsymbol{U}} \Big[\mathcal{L}^u V_\alpha(x) + c(x,u) \Big] = \alpha V_\alpha(x) , \tag{2.6}$$

and the α-discounted optimal policy v_α^* realizes the pointwise infimum in (2.6).

Under the near-monotone assumption and Lipschitz condition on $c(x,u)$, the optimal policy for the average cost criterion can be characterized as follows: there exists a unique (V, ρ) satisfying

$$\begin{gathered} \inf_{u \in \boldsymbol{U}} \Big[\mathcal{L}^u V(x) + c(x,u) \Big] = \rho \\ \rho \leq \rho^* , \quad V(0) = 0 , \quad \inf_{\mathbb{R}^n} V > -\infty , \end{gathered}$$

and the optimal control is the minimizer of the above, and is a stable stationary Markov control.

Chapter 3

Optimal Sensor Querying: General Markovian Models with Controlled Observations

3.1 Motivation

In recent years, much attention has been paid to networked control systems (NCS), in which the sensors, the controllers and the actuators are located in a distributed manner and are interconnected by communication channels. In such systems, the information collected by sensors and the decisions made by controllers are not instantly available to the controllers and actuators, respectively. Rather they are transmitted through communication channels, which might suffer delay and/or transmission errors, and as such this transmission carries a cost. Understanding the interaction between the control system and the communication system becomes more and more important and plays a key role on the overall performance of NCS.

One simple example of NCS is an automobile manufacturing system shown in Figure 1.2: there are a large amount of sensors deployed in the system for sensing, detection and data collection; data collected by sensors is transmitted to the processing center through communication channels; the processing center (or decision maker) makes decisions according to received information and transmits decisions to robots (actuators) through communication channels.

However, as the number of sensors in the network becomes larger and larger, it is prohibitive to gather data from all the sensors at a time. Thus the processing center has to make decisions on which sensors to be queried at a given time according to some optimization criteria.

Broadly speaking, the amount of information the controller receives, affects the performance of estimation and control. The more information the processing center can obtain, the better performance it can achieve. However, information is not free. On the one hand, it consumes resources such as bandwidth, and power (i.e., in the case of a wireless channel), while on the other, by generating more traffic in the network it induces delays. If one incorporates in a standard optimal control problem an additional running penalty, associated with receiving the observations at the controller, then a tradeoff would result that balances the cost of observation and the performance of the control. We consider a simple network scenario: a network of sensors, provides observation on the system state sent to the controller through a communication channel. The controller has the option of requesting different amounts of information from the sensors (i.e., more detailed or coarser observations), and can do so at each discrete time step. Based on the information requested, an estimate of the state is computed and a control action is decided upon. However, what is different here is that there is a running cost, associated with the information requested, which is added to the running cost of the original control criterion. As a result the observation space is not static, but rather changes dynamically as the controller issues different queries on the sensors.

In our work, the optimal sensor querying problem is studied in the context of partially observable Markov decision processes (POMDP). In this chapter, we formulate it as a general POMDP problem and specialize to a class of problems with hierarchical observations. In the next chapter, we focus on a special class of Markovian systems – stochastic linear systems, in which some important structural results can be obtained.

This chapter is organized as follows. We first introduce a general POMDP model enabled with sensor querying in Section 3.2, and then discuss the background of the topic and the related work in Section 3.3. Finally some structural results are obtained for a class of the problems with hierarchical observations.

3.2 System model: POMDPs with controlled observations

We consider the control of a dynamical system (shown in Figure 3.1), which is governed by a Markov chain $(\boldsymbol{X}, \boldsymbol{U}, P, \mu)$, where $\boldsymbol{X}$ is the state space (assumed to be a Borel space), μ is the initial distribution of the state variable X_t and $\boldsymbol{U}$ is the set of actions, which is assumed to be a compact metric space. We use capital letters to denote random processes and variables and lower case letters to denote the elements of a space. We denote by $\mathcal{P}(\boldsymbol{X})$ the set of probability measures on $\boldsymbol{X}$. The dynamics of the process are governed by a transition kernel P on $\boldsymbol{X}$ given $\boldsymbol{X} \times \boldsymbol{U}$, which may be interpreted as

$$P_u(A \mid x) = \mathrm{Prob}(X_{t+1} \in A \mid X_t = x,\ U_t = u)\,,$$

for $t = 0, 1, \ldots,$ and A an element of the set of the Borel σ-field of $\boldsymbol{X}$, the latter denoted by $\mathfrak{B}(\boldsymbol{X})$.

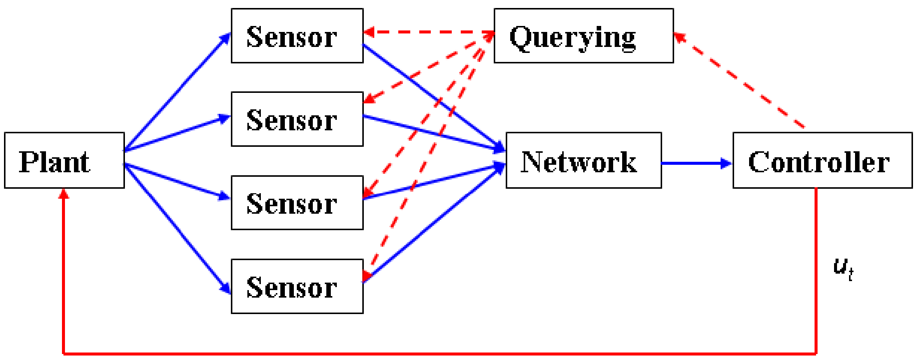


Figure 3.1: The system block of a networked control system with sensor querying

The model includes ℓ distinct observation processes, but only one of these can be accessed at a time. Consider for example, a network of sensors providing observations for the control of a dynamical system. Suppose that there are ℓ levels of sensor information, and at each time t, Y_t^i represents the set of data provided at the i-th level, which lives in a space $\boldsymbol{Y}^{(i)}$. In as much as the complete set of data is a partial measurement of the state X_t of the system, we are provided with stochastic kernels $\mathcal{K}^i$ on $\mathcal{P}(\boldsymbol{Y}^{(i)})$ given $\boldsymbol{X}$, which may be interpreted as the conditional distribution of Y_t^i given X_t, i.e.,

$$\mathcal{K}^i(y \mid x) = \mathrm{Prob}(Y_t^i = y \mid X_t = x)\,.$$

The mechanism of sensor querying is facilitated by the query variable Q_t which chooses the subset of sensors to be queried at time t, i.e., takes values in $\boldsymbol{Q} = \{1, \ldots, \ell\}$. The evolution of the system is as follows: at each time t an action and query $(U_t, Q_t) = (u, q) \in \boldsymbol{U} \times \boldsymbol{Q}$ are chosen and the system moves to the next state X_{t+1} according to the probability transition function P_u, and the data set $Y_{t+1}^q \in \boldsymbol{Y}^{(q)}$, corresponding to the queried sensors, is obtained. As shown in Figure 3.1, the blue arrows indicate the data path of the system and the red arrows denote the control path including the control of plant U_t and the control of querying Q_t.

One special case of this model is when the levels of sensor information constitute a hierarchy, i.e., the data set becomes richer as we move up in the levels, meaning that the σ-fields are ordered by the inclusion $\sigma(Y_t^1) \subset \cdots \subset \sigma(Y_t^\ell)$. Another scenario, in the sensor scheduling problem, involves ℓ independent sensors with observations Y_t^i, and at each time t, only one can be accessed (e.g., due to interference).

Following the standard POMDP model formulation (e.g., [33]), we define $\boldsymbol{Y} \triangleq \bigcup_{q \in \boldsymbol{Q}} \boldsymbol{Y}^{(q)}$, and the history spaces $\{\boldsymbol{H}_t\}$ by $\boldsymbol{H}_0 \triangleq \mathcal{P}(\boldsymbol{X})$ and

$$\boldsymbol{H}_{t+1} \triangleq \boldsymbol{H}_t \times \boldsymbol{U} \times \boldsymbol{Q} \times \boldsymbol{Y}\,, \quad t = 0, 1, \ldots\,.$$

The information available for decision making at time t is the history $\mathcal{H}_t = \sigma\{H_t\}$, where $\{H_t\}$ stands for the history process.

An *admissible* control is a sequence $v = (v_0, v_1, \ldots)$, where each v_t is a kernel on $\boldsymbol{U} \times \boldsymbol{Q}$ given $\boldsymbol{H}_t$. Specifying an initial distribution μ and an admissible control v, determines a unique probability measure $\mathbb{P}_\mu^v$ on the path space of the process. such that for $A \in \mathfrak{B}(\boldsymbol{X})$, $C \in \mathfrak{B}(\boldsymbol{U} \times \boldsymbol{Q})$, and $D \in \mathfrak{B}(\boldsymbol{Y})$, the following hold $\mathbb{P}^v_\mu$–a.s.

$$
\begin{aligned}
\mathbb{P}^v_\mu(X_0 \in D) &= \mu(D) \\
\mathbb{P}^v_\mu\big((U_t \times Q_t) \in C\big) &= v_t(C \mid H_t) \\
\mathbb{P}^v_\mu(X_{t+1} \in A \mid H_t, X_t) &= P_{U_t}(D \mid X_t) \\
\mathbb{P}^v_\mu(Y_{t+1} \in D \mid H_t, X_t) &= \int_{\boldsymbol{X}\times\boldsymbol{U}\times\boldsymbol{Q}} \mathcal{K}^q(D \mid x) P_u(dx \mid X_t) v_t(du \times dq \mid H_t)\,.
\end{aligned}
$$

Markov controls and stationary controls are defined in the standard manner. We let $\mathcal{V}$ denote all admissible controls, and $\mathcal{V}_M$, $\mathcal{V}_S$ all the Markov, stationary (Markov) controls respectively. Under a Markov control v the probability measure $\mathbb{P}^v_\mu$ renders (X_t, Y_t) a Markov process.

Following the theory of partially observed stochastic systems, we obtain an equivalent completely observed model through the introduction of the conditional distribution Ψ_t of the state given the observations [3, 33, 35]. The process Ψ_t lives in $\boldsymbol{\Psi} \triangleq \mathcal{P}(\boldsymbol{X})$. An important difference from the otherwise routine construction is that the observation process does not live in a fixed space but varies dynamically based on the query process. The query variable selects the observation space, and the nonlinear Bayesian filter that updates the state estimate is chosen accordingly. Let

$$
\begin{aligned}
\tilde{P}(dx, dy \mid \psi, u, q) &\triangleq \int_{x' \in \boldsymbol{X}} \mathcal{K}^q(dy|x) P_u(dx \mid x')\psi(dx') \\
V(dy, \psi, u, q) &\triangleq \int_{x \in \boldsymbol{X}} \tilde{P}(dx, dy \mid \psi, u, q)\,.
\end{aligned}
$$

Decomposing the measure $\tilde{P}$ as

$$
\tilde{P}(dx, dy \mid \psi, u, q) = \mathcal{T}(\psi, y, u, q)(dx)\, V(dy, \psi, u, q)\,,
$$

we obtain the filtering equation

$$
\psi_{t+1} = \mathcal{T}(\psi_t, y_{t+1}, u_t, q_t)(dx)\,. \tag{3.1}
$$

The nonlinear filtering operator $\mathcal{T}$ that has the intuitive interpretation of the a posteriori conditional distribution of the state, given that decision (u, q) was made, observation $y \in \boldsymbol{Y}^{(q)}$ obtained, and an a priori distribution ψ. Likewise, $V(dy, \psi, u, q)$ is interpreted as the one-step ahead conditional probability on the observation space $\boldsymbol{Y}^{(q)}$ given an a priori distribution ψ for the state, under decision u.

The model includes a running penalty $r : \boldsymbol{X} \times \boldsymbol{U} \to \mathbb{R}$, which is assumed to be continuous and non-negative, as well as a penalty function $c : \boldsymbol{Q} \to \mathbb{R}$ that represents the cost of information. Let $g = r + c$. We are interested primarily in the long-term average, or ergodic criterion. In other words, we seek to minimize, over all admissible policies $v \in \mathcal{V}$,

$$J^v \triangleq \limsup_{N\to\infty} \frac{1}{N} \mathbb{E}^v_x \left[\sum_{t=0}^{N-1} g(X_t, U_t, Q_t) \right] . \tag{3.2}$$

When the expectation operator is omitted in (3.2), J^v is referred to as the ergodic criterion. We also consider the β-discounted criterion

$$J^v_\beta(x) \triangleq \mathbb{E}^v_x \left[\sum_{t=0}^{N-1} \beta^t g(X_t, U_t, Q_t) \right] , \quad \beta \in (0, 1) . \tag{3.3}$$

We define

$$J^* \triangleq \inf_{v\in\mathcal{V}} J^v , \qquad J^*_\beta(x) \triangleq \inf_{v\in\mathcal{V}} J^v_\beta(x) .$$

If we let $\tilde{g}(\psi, u, q) = \int g(x, a, q)\psi(dx)$, the control criteria in (3.2)–(3.3) can be expressed in the equivalent CO model.

Stationary optimal policies for the β-discounted cost objective can be characterized via the HJB equation,

$$J_\beta(\psi) = \min_{(u,q)\in\boldsymbol{U}\times\boldsymbol{Q}} \left\{ \tilde{g}(\psi, u, q) + \beta \int_{\boldsymbol{Y}^{(q)}} V(dy, \psi, u, q) J_\beta\big(\mathcal{T}(y, \psi, u, q)\big) \right\} , \tag{3.4}$$

where J_β is the optimal value function. For the long-term average or ergodic objective, the HJB equation takes the form

$$J^* + h(\psi) = \min_{(u,q)\in\boldsymbol{U}\times\boldsymbol{Q}} \left\{ \tilde{g}(\psi, u, q) + \int_{\boldsymbol{Y}^{(q)}} V(dy, \psi, u, q) h\big(\mathcal{T}(y, \psi, u, q)\big) \right\} . \tag{3.5}$$

In (3.5), J^* is the optimal average cost, and h is called the bias function.

3.3 Background and related work

Early work on the control of the observation process (or measurement) can be traced back to the seminal paper of Meier *et al* [71]. The separation principle between the optimal plant control and optimal measurement control was proved for finite-horizon linear quadratic Gaussian (LQG) control. With this property, the optimal plant control can be designed independently of the measurement control, which can significantly simplify the synthesis of the optimal controller.

Later on, work has focused on the optimal measurement control, or the so-called sensor scheduling problem [7, 8, 48, 57, 70, 72, 73, 90, 92], in which there are a number of sensors with different levels of precision and operation costs and the scheduler can access only one sensor at a time to receive the observation. The objective is to minimize a weighted average of the estimation error and observation cost.

The problem is studied in different Markovian systems. For example, in [90] and [72], the sensor scheduling problem is addressed for continuous-time linear systems; while in [57], the system dynamics corresponds to a general hidden Markov chain. With little assumption on the model as [57], the average cost optimal control can only be characterized by (3.5), which yields little structural results and does not even guarantee the existence of the optimal ergodic control with a constant J^* independent of the initial condition. Some approximation algorithms are proposed as well in [57] to obtain the optimal solution via numerical methods.

On the other hand, a lot of attention has been paid to the stochastic linear model, e.g., [7, 42, 70, 71], in which some much simpler forms can be obtained thanks to the nice linear structure of the system dynamics. Recently in [42], Gupta *et al* propose computationally tractable algorithms to solve the stochastic sensor scheduling problem for the finite-horizon LQG problem.

3.4 POMDPs with hierarchical observations

In this section we focus on models with the hierarchical structure $\sigma(Y_t^1) \subset \cdots \subset \sigma(Y_t^\ell)$. We consider a POMDP model with finite state space $\boldsymbol{X} = \{1, \ldots, n\}$ and observation spaces $\boldsymbol{Y}^{(q)} = \{1, \ldots, l_q\}$, $q \in \boldsymbol{Q}$. The action space $\boldsymbol{U}$ is assumed to be a compact metric space. The dynamics of the process are governed by a transition kernel on $\boldsymbol{X} \times \boldsymbol{Y}$,

$$Q^q_{ij}(y,u) = \mathbb{P}(X_{t+1} = j, Y^q_{t+1} = y \mid X_t = i, U_t = u)\,.$$

For fixed q, y and u, Q^q can be viewed as an $n \times n$ substochastic matrix and is assumed continuous with respect to u. Representing ψ as a row vector of dimension n, (3.1) takes the form

$$\begin{aligned} V(\psi,y,u,q) &= \psi Q^q(y,u)\mathbf{1} \\ \mathcal{T}(\psi,y,u,q) &= \begin{cases} \frac{\psi Q^q(y,u)}{V(\psi,y,u,q)} & \text{if } \; V(\psi,y,u,q) > 0 \\ \bar{\psi} & \text{otherwise}\,, \end{cases} \end{aligned}$$

where $\bar{\psi}$ can be chosen arbitrarily.

Under the hierarchical structure assumed, the observation space $\boldsymbol{Y}^{(q+1)}$, admits a partition $\{S^q_y\,,\; y \in \boldsymbol{Y}^{(q)}\}$, satisfying the property

$$Q^q(y,u) = \sum_{y' \in S^q_y} Q^{q+1}(y',u)\,, \quad \forall (y,u) \in \boldsymbol{Y}^{(q)} \times \boldsymbol{U}\,. \tag{3.6}$$

Note that (3.6) implies that for any $y \in \boldsymbol{Y}^{(q)}$, $\mathcal{T}(\psi,y,u,q)$ can be expressed as convex combination of $\{\mathcal{T}(\psi,y',u,q+1)\,,\; y' \in S^q_y\}$.

Next we employ results from [50] to show existence of a solution to (3.5) for a POMDP with hierarchical observations, by imposing a condition on the (finest) observation space $\boldsymbol{Y}^{(\ell)}$.

Next, we use some of results in [50] to prove existence of a solution to (3.5) with hierarchical observations, by imposing a condition only on the (finest) observation space $\boldsymbol{Y}^{(\ell)}$.

Remark 3.4.1. *According to the results in [50], under Assumption 2.2.2, there exists a solution to* (3.5)*, provided the observation space is restricted to* $\boldsymbol{Y}^{(\ell)}$.

We have the following existence theorem.

Theorem 3.4.1. *Let Assumption 2.2.2 hold. Then, there exists a solution* (J^*, h) *to* (3.5)*, with* $J^* \in \mathbb{R}$ *and* $h : \boldsymbol{\Psi} \to \mathbb{R}$*, a concave function. Moreover, the minimizer* *in* (3.5) *defines a stationary optimal policy relative to the ergodic criterion, and* J^* *is the optimal cost.*

Proof. By (3.6), for each $q \in \boldsymbol{Q}$, there exists a partition $\{\tilde{S}^q_y\,,\ y \in \boldsymbol{Y}^{(q)}\}$ of $\boldsymbol{Y}^{(\ell)}$, satisfying

$$Q^q(y,u) = \sum_{y' \in \tilde{S}^q_y} Q^\ell(y',u)\,, \quad \forall (y,u) \in \boldsymbol{Y}^{(q)} \times \boldsymbol{U}\,. \tag{3.7}$$

Then for any $q^k \in \boldsymbol{Q}^k$, $y^k \in \boldsymbol{Y}^{(q^k)}$, and $u^k \in \boldsymbol{U}^k$, we have

$$\begin{aligned} Q(y^k,u^k,q^k) &= \prod_{t=0}^{k-1} Q^{q_t}(y_{t+1},u_t) \\ &= \prod_{t=0}^{k-1} \sum_{y' \in \tilde{S}^{q_t}_{y_{t+1}}} Q^\ell(y',u_t) \\ &= \sum_{\hat{y}^k \in \tilde{S}^{q_0}_{y_1} \times \cdots \times \tilde{S}^{q_{k-1}}_{y_k}} Q(\hat{y}^k,u^k,\{\ell\}^k)\,. \end{aligned} \tag{3.8}$$

Assuming (2.3) holds, fix $i \in \boldsymbol{X}$, $u^{k_0} = \{u_0,\dots,u_{k_0-1}\}$ and $q^{k_0} = \{q_0,\dots,q_{k_0-1}\}$, and let $\bar{y}^k$, $k \le k_0$ be such that

$$Q_{ij}(\bar{y}^k,u^k,\{\ell\}^k) > 0\,, \quad \forall j \in \boldsymbol{X}\,. \tag{3.9}$$

Since $\{\tilde{S}^q_y\,,\ y \in \boldsymbol{Y}^{(q)}\}$ is a partition of $\boldsymbol{Y}^{(\ell)}$, choose $y_t \in \boldsymbol{Y}^{(q_t)}$ such that $\bar{y}_t \in \tilde{S}^{q_t}_{y_t}$, for all $t = 0,\dots,k$. By (3.8)–(3.9),

$$Q_{ij}(y^k,u^k,q^k) \ge Q_{ij}(\bar{y}^k,u^k,\{\ell\}^k) > 0\,, \quad \forall j \in \boldsymbol{X}\,.$$

Therefore, Assumption 2.2.2 (Assumption 4 in [50]) is satisfied, which yields the result. $\square$

Chapter 4

Optimal Sensor Querying: LQG Models

4.1 Introduction

In the last chapter, we have formulated the sensor querying problem as a POMDP with controlled observation, in which the state space is finite. With additional assumption on the information structure of the observation spaces, the existence of ergodic control has been shown under mild conditions.

In this chapter, we consider the sensor querying problem in a stochastic linear system, in which the state space is infinite. We consider the LQG control problem and prove that a partial separation principle of estimation and control holds over the infinite horizon: the optimal control can be decoupled into two subproblems, an optimal control problem with full observations and an optimal query/estimation problem requiring the knowledge of the controller gain. The estimation problem reduces to a Kalman filter, with the gain computed by a discrete algebraic Riccati equation (DARE). However, the optimal query is characterized by a dynamic programming equation.

The rest of this chapter is organized as follows. In Section 4.2 we introduce the LQG control model with sensor querying and list some important results in the traditional LQG control model. Section refsec:LQG-T shows the partial separation principle for the optimal control in the finite horizon. In Section4.4, we show the necessary and sufficient condition for stabilizability issues. In Section4.5, the infinite-horizon optimal control problem, including the discounted and average costs, is studied and the dynamic programming equation is further simplified and decoupled into two separate problems: (a) optimal estimation problem, and (b) control, the latter being a standard LQG optimal control problem. In Section 4.6 we present some examples.

4.2 Linear quadratic Gaussian (LQG) control: the model and main results

4.2.1 LQG model

Consider a linear system governed by,

$$X_{t+1} = AX_t + BU_t + DW_t\,, \quad t = 0, 1, \ldots, \tag{4.1}$$

where $X_t \in \mathbb{R}^{N_x}$ is the system state, $U_t \in \mathbb{R}^{N_u}$ is the control, and the noise process $\{W_t\}$ is i.i.d., and normally distributed. We assume that X_0 is Gaussian with mean $\bar{x}_0$ and covariance matrix Σ_0, and denote this by $X_0 \sim \mathcal{N}(\bar{x}_0, \Sigma_0)$. We also assume that X_0 and $\{W_t\,,\ t \geq 0\}$ are independent. The state-process is being observed by

$$Y_t = C_{Q_{t-1}} X_t + FW_t\,, \quad t \geq 1\,, \tag{4.2}$$

with $Y_t \in \mathbb{R}^{N_y}$, and $\det(FF^{\mathsf{T}}) \neq 0$. Moreover, we assume the system noise and observation noise are independent, i.e., $DF^{\mathsf{T}} = 0$. This independence assumption results in a simplification of the algebra; otherwise, it is not essential.

The running cost r is quadratic in the state and control, and takes the form

$$r(x, u) = x^{\mathsf{T}} Rx + u^{\mathsf{T}} Su\,,$$

where R and S belong to $\mathcal{M}^+$, the set of symmetric, positive definite matrices in $\mathbb{R}^{N_x \times N_x}$.

4.2.2 Main results in LQG control

Optimal control for the stochastic linear system (4.1)–(4.2) with fixed query has well known results, which we are going to review briefly in this section.

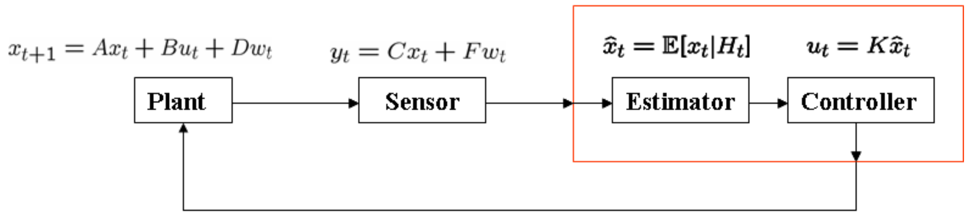


Figure 4.1: The optimal controller for LQG problem: separation principle

For finite-horizon cost with quadratic penalty on the control and the system state, it has been shown that a separation principle holds, namely, the optimal control rule given the system observations can be decomposed into two components (see Figure 4.1): one component is *estimator*, which updates the estimates of the current system state, and another component is *controller*, which is a linear feedback control based on the output of the estimator. Moreover, the design of the controller and estimator is independent to each other, thus they can be synthesized independently. For the estimator, *Kalman filter* has been shown to be optimal, and the mean $\hat{X}_t = \mathbb{E}[X_t \mid \mathcal{Y}^t]$ and error variance $\hat{\Pi}_{t+1} = \mathbb{E}[(X_t - \hat{X}_t)^2]$ are updated as follows,

$$\begin{aligned}
\hat{X}_{t+1} &= A\hat{X}_t + BU_t + \hat{K}_{t+1}\big(Y_{t+1} - C(A\hat{X}_t + BU_t)\big) \\
\hat{K}_{t+1} &= \Xi(\hat{\Pi}_t)C^\top (C\Xi(\hat{\Pi}_t)C^\top + FF^\top)^{-1} \\
\hat{\Pi}_{t+1} &= \Xi(\hat{\Pi}_t) - \Xi(\hat{\Pi}_t)C^\top (C\Xi(\hat{\Pi}_t)C + FF^\top)^{-1} C\Xi(\hat{\Pi}_t) \\
\Xi(\hat{\Pi}_t) &\triangleq DD^\top + A\hat{\Pi}_t A^\top .
\end{aligned}$$

Note here $\hat{\Pi}_t$ is determined by forward recursion. For the control, linear feedback control is optimal, namely, the optimal solution is

$$U_t^* = -K_t \hat{X}_t ,$$

where the feedback gain is determined backward recursively by

$$\begin{aligned}
K_t &= (B^\top \Pi_{t+1} B + S)^{-1} B^\top \Pi_{t+1} A \\
\Pi_t &= R + A^\top \Pi_{t+1} A - A^\top \Pi_{t+1} B (B^\top \Pi_{t+1} B + S)^{-1} B^\top \Pi_{t+1} A .
\end{aligned}$$

For the infinite-horizon problem, under the conditions (A, B) is stabilizable, $(R^{\frac{1}{2}}, A)$ is detectable, the backward recursion for the controller converges, namely $\lim_{t\to\infty} \Pi_t = \Pi$ and Π is the unique positive definite solution for a discrete algebraic Riccati equation (DARE). On the other hand, under the conditions (C, A) is detectable and (A, D) is stabilizable, the forward recursion for the estimator converges, namely, $\lim_{t\to\infty} \hat{\Pi}_t = \hat{\Pi}$ and $\hat{\Pi}$ is the unique positive definite solution for another DARE.

4.3 Optimal control over a finite horizon

For an initial condition X_0 and an admissible policy $v = \{(U_t, Q_t),\ t \geq 0\}$, let $\mathbb{P}^v$ denote the unique probability measure on the pathspace of the process, and $\mathbb{E}^v$ the corresponding expectation operator. Whenever needed we indicate the dependence on X_0 explicitly (or more precisely the dependence on the law of X_0), by using the notation $\mathbb{P}^v_{X_0}$ and $\mathbb{E}^v_{X_0}$. The optimal control problem over a finite horizon N, amounts to minimizing over all admissible controls the functional

$$J^v_N \triangleq \mathbb{E}^v \left[\sum_{t=0}^{N-1} \big(c(Q_t) + r(X_t, U_t)\big) + X_N^{\mathsf{T}} \Pi_N X_N \right], \tag{4.3}$$

where $\Pi_N \in \mathcal{M}_0^+$, the set of symmetric, positive semi-definite matrices in $\mathbb{R}^{N_x \times N_x}$. In (4.3), J^v_N is of course a function of the law of X_0, and hence can be parameterized as $J^v_N = J^v_N(\bar{x}_0, \Sigma_0)$.

Theorem 4.3.1. *Consider the control system in* (4.1)–(4.2)*, under the assumptions stated in Section 4.2, and let* $\bar{x}_0$ *and* Σ_0*, be the mean and covariance matrix of* X_0*, respectively. Let*

$$J^*_N(\bar{x}_0, \Sigma_0) \triangleq \inf_{v \in \mathcal{V}} J^v_N(\bar{x}_0, \Sigma_0)\,.$$

Let $v^* = \{(U^*_t, Q^*_t)\,,\ t = 0, \dots, N-1\}$*, where* $\{U^*_t\}$ *is defined by*

$$U^*_t = -K_t \hat{X}_t\,, \tag{4.4}$$

where

$$K_t = (B^{\mathsf{T}} \Pi_{t+1} B + S)^{-1} B^{\mathsf{T}} \Pi_{t+1} A \tag{4.5a}$$

$$\varPi_t = R + A^\mathsf{T} \varPi_{t+1} A - A^\mathsf{T} \varPi_{t+1} B (B^\mathsf{T} \varPi_{t+1} B + S)^{-1} B^\mathsf{T} \varPi_{t+1} A \,, \tag{4.5b}$$

and Q_t^* is a selector of the minimizer in the dynamic programming equation

$$f_t(\hat{\varPi}) = \min_q \big\{ c(q) + \operatorname{tr}(\tilde{\varPi}_t \hat{\varPi}) + f_{t+1}(\mathcal{T}_q(\hat{\varPi})) \big\} \,, \qquad t = 0, \dots, N-1 \,, \tag{4.6}$$

with $f_N \triangleq 0$, where

$$\tilde{\varPi}_t \triangleq R - \varPi_t + A^\mathsf{T} \varPi_{t+1} A \,, \quad t = 0, \dots, N-1 \,.$$

Then v^* is optimal with respect to the cost functional J_N^v and

$$J_N^*(\bar{x}_0, \varSigma_0) = \tilde{J}_N^{U^*}(\bar{x}_0, \varSigma_0) + f_0(\varSigma_0) \,.$$

Proof. Let $Y^t = \{Y_1, \dots, Y_t\}$, and $\mathcal{Y}^t = \sigma(Y^t)$. Invoking the results of the general POMDP model in Section 3.2, we can obtain an equivalent completely observed model using the conditional distribution of X_t given $\mathcal{Y}^t$ as the new state. It is well known that with respect to $\mathbb{P}^v$ the conditional distribution of X_t given $\mathcal{Y}^t$ is Gaussian [12]. Let $\hat{X}_t = \mathbb{E}^v[X_t \mid \mathcal{Y}^t]$. Since there is no observation Y_0 in our model, we set $\mathcal{Y}^0$ as the trivial σ-field. Hence, $\hat{X}_0 = \mathbb{E}[X_0] = \bar{x}_0$. Then, a standard derivation, yields

$$\hat{X}_{t+1} = A\hat{X}_t + BU_t + \hat{K}_{t+1}\big(Y_{t+1} - C_{Q_t}(A\hat{X}_t + BU_t)\big) \,, \tag{4.7}$$

where

$$\hat{K}_{t+1} = \varXi(\hat{\varPi}_t) C_{Q_t}^\mathsf{T} (C_{Q_t} \varXi(\hat{\varPi}_t) C_{Q_t}^\mathsf{T} + FF^\mathsf{T})^{-1} \tag{4.8a}$$

$$\hat{\varPi}_{t+1} = \varXi(\hat{\varPi}_t) - \varXi(\hat{\varPi}_t) C_{Q_t}^\mathsf{T} (C_{Q_t} \varXi(\hat{\varPi}_t) C_{Q_t} + FF^\mathsf{T})^{-1} C_{Q_t} \varXi(\hat{\varPi}_t) \tag{4.8b}$$

$$\varXi(\hat{\varPi}_t) \triangleq DD^\mathsf{T} + A\hat{\varPi}_t A^\mathsf{T} \,. \tag{4.8c}$$

In (4.8), $\hat{\varPi}_t$ is the conditional covariance of $X_t - \hat{X}_t$ under $\mathbb{P}^v$ given $\mathcal{Y}^t$, and $\hat{\varPi}_0 = \varSigma_0$. By (4.8b)–(4.8c), the conditional error covariance matrix $\hat{\varPi}_t$ satisfies $\hat{\varPi}_{t+1} =$

$\mathcal{T}_{Q_t}(\hat{\Pi}_t)$, where

$$\mathcal{T}_q(\hat{\Pi}) \triangleq \Xi(\hat{\Pi}) - \Xi(\hat{\Pi})C_q^{\mathsf{T}}(C_q\Xi(\hat{\Pi})C_q^{\mathsf{T}} + FF^{\mathsf{T}})^{-1}C_q\Xi(\hat{\Pi})\,. \tag{4.9}$$

If an admissible sequence $\{Q_t,\ t \geq 0\}$ is specified, then standard LQG theory shows that the policy $\{U_t^*,\ t = 0, \dots, N-1\}$, given by (4.4)–(4.5), is optimal relative to the functional J_N^v. In other words, if we denote $J_N^v = J_N^{U,Q}$ with $Q = \{Q_t,\ t \geq 0\}$ fixed, then $U^* = \{U_t^*,\ t \geq 0\}$ defined by (4.4) satisfies

$$J_N^{U^*,Q} = \inf_{\tilde{U}}\ J_N^{\tilde{U},Q}\,,$$

where the infimum is over all admissible policies $\tilde{U}$.

Combining the feedback policy in (4.4) with (4.7), we obtain

$$\hat{X}_{t+1} = (A - BK_t)\hat{X}_t + \hat{K}_{t+1}C_{Q_t}\big(A(X_t - \hat{X}_t) + DW_t)\big) + \hat{K}_{t+1}FW_{t+1}\,. \tag{4.10}$$

A straightforward computation using (4.5)–(4.10), yields

$$\begin{aligned}\mathbb{E}^{U^*,Q}[X_t^{\mathsf{T}}\Pi_tX_t] &= \mathbb{E}^{U^*,Q}[\hat{X}_t^{\mathsf{T}}\Pi_t\hat{X}_t] + \mathbb{E}^{U^*,Q}[(X_t - \hat{X}_t)^{\mathsf{T}}\Pi_t(X_t - \hat{X}_t)]\\ &= \mathbb{E}^{U^*,Q}[\hat{X}_{t-1}^{\mathsf{T}}(\Pi_{t-1} - R - K_{t-1}^{\mathsf{T}}SK_{t-1})\hat{X}_{t-1}]\\ &\qquad + \mathbb{E}^{U^*,Q}[\operatorname{tr}(\Pi_t\hat{\Pi}_t) + \operatorname{tr}(\Pi_t\tilde{\hat{\Pi}}_t)]\,, \end{aligned} \tag{4.11}$$

where

$$\tilde{\hat{\Pi}}_{t+1} \triangleq \Xi(\hat{\Pi}_t)C_{Q_t}^{\mathsf{T}}(C_{Q_t}\Xi(\hat{\Pi}_t)C_{Q_t}^{\mathsf{T}} + FF^{\mathsf{T}})^{-1}C_{Q_t}\Xi(\hat{\Pi}_t)\,.$$

Similarly, for $t = 0, \dots, N-1$, we have

$$\mathbb{E}^{U^*,Q}[r(X_t, U_t^*)] = \mathbb{E}^{U^*,Q}\big[\hat{X}_t^{\mathsf{T}}(R + K_t^{\mathsf{T}}SK_t)\hat{X}_t + \operatorname{tr}(R\hat{\Pi}_t)\big]\,. \tag{4.12}$$

Thus by (4.11)–(4.12), for $t = 0, \dots, N-1$,

$$\begin{aligned}&\mathbb{E}^{U^*,Q}[r(X_t, U_t^*) + X_{t+1}^{\mathsf{T}}\Pi_{t+1}X_{t+1}]\\ &\quad = \mathbb{E}^{U^*,Q}[\hat{X}_t^{\mathsf{T}}\Pi_t\hat{X}_t + \operatorname{tr}(\Pi_{t+1}\hat{\Pi}_{t+1}) + \operatorname{tr}(\Pi_{t+1}\tilde{\hat{\Pi}}_{t+1})] + \mathbb{E}^{U^*,Q}[\operatorname{tr}(R\hat{\Pi}_t)]\,. \end{aligned} \tag{4.13}$$

Since

$$\operatorname{tr}(R\hat{\mathit{\Pi}}_t) = \operatorname{tr}(\tilde{\mathit{\Pi}}_t\hat{\mathit{\Pi}}_t) - \operatorname{tr}(A^{\mathsf{T}}\mathit{\Pi}_{t+1}A\hat{\mathit{\Pi}}_t) + \operatorname{tr}(\mathit{\Pi}_t\hat{\mathit{\Pi}}_t)\,, \tag{4.14}$$

and

$$\operatorname{tr}(\mathit{\Pi}_t\hat{\mathit{\Pi}}_t) + \operatorname{tr}(\mathit{\Pi}_t\tilde{\hat{\mathit{\Pi}}}_t) = \operatorname{tr}(\mathit{\Pi}_t DD^{\mathsf{T}}) + \operatorname{tr}(\mathit{\Pi}_t A\hat{\mathit{\Pi}}_{t-1}A^{\mathsf{T}})\,, \tag{4.15}$$

we have,

$$\operatorname{tr}(R\hat{\mathit{\Pi}}_{t-1}) + \operatorname{tr}(\mathit{\Pi}_t\hat{\mathit{\Pi}}_t) + \operatorname{tr}(\mathit{\Pi}_t\tilde{\hat{\mathit{\Pi}}}_t) = \operatorname{tr}(\tilde{\mathit{\Pi}}_{t-1}\hat{\mathit{\Pi}}_{t-1}) + \operatorname{tr}(\mathit{\Pi}_{t-1}\hat{\mathit{\Pi}}_{t-1}) + \operatorname{tr}(\mathit{\Pi}_t DD^{\mathsf{T}})\,. \tag{4.16}$$

Define

$$J^v_{t,N} \triangleq \mathbb{E}^v\left[\sum_{k=t}^{N-1}\big(c(Q_k) + r(X_k,U_k)\big) + X_N^{\mathsf{T}}\mathit{\Pi}_N X_N\right].$$

Simple induction, using (4.13) and (4.16) yields, for $t = 0,\ldots,N$,

$$J^{U^*,Q}_{t,N} = \mathbb{E}^{U^*,Q}[\hat{X}_t^{\mathsf{T}}\mathit{\Pi}_t\hat{X}_t + \operatorname{tr}(\mathit{\Pi}_t\hat{\mathit{\Pi}}_t)] + \mathbb{E}^{U^*,Q}\left[\sum_{k=t}^{N-1}\big(c(Q_k) + \operatorname{tr}(\tilde{\mathit{\Pi}}_k\hat{\mathit{\Pi}}_k) + \operatorname{tr}(\mathit{\Pi}_{k+1}DD^{\mathsf{T}})\big)\right].$$

Therefore, $J^{U^*,Q}_N = \tilde{J}^{U^*}_N + \hat{J}^{U^*,Q}_N$, where

$$\tilde{J}^{U^*}_N \triangleq \bar{x}_0^{\mathsf{T}}\mathit{\Pi}_0\bar{x}_0 + \operatorname{tr}(\mathit{\Pi}_0\mathit{\Sigma}_0) + \sum_{k=1}^{N}\operatorname{tr}(\mathit{\Pi}_k DD^{\mathsf{T}})\,,$$

which does not depend on Q, and

$$\hat{J}^{U^*,Q}_N \triangleq \mathbb{E}^{U^*,Q}\left[\sum_{k=0}^{N-1}\big(c(Q_k) + \operatorname{tr}(\tilde{\mathit{\Pi}}_k\hat{\mathit{\Pi}}_k)\big)\right]. \tag{4.17}$$

Define $f_t(\hat{\mathit{\Pi}})$ as the cost-to-go function for (4.17), then the optimal policy Q^* can be determined by (4.6) according to the dynamic programming principle. □

As in the standard theory of LQG control with partial observations, the optimal control of (4.1)–(4.2) is a *certainty equivalence* control, namely, the optimization problem can be separated into two stages: first, the optimal control U_t^* is the linear feedback control in (4.4) whose gain does not depend on the choice of the query policy $\{Q_t\}$; second, the conditional distribution of the system state is obtained recursively via the filtering equation (4.7) which couples with the dynamic programming equation (4.6) to determine the optimal query policy. The difference from the standard LQG problem is that the dynamic programming equation depends on the controller gain which evolves according to (4.5). Thus, (4.6) can be viewed as the solution of an optimal estimation problem, in which the cost function is the sum of the cost of the query and a weighted estimation error (4.17).

4.4 Stabilization

Stability is a basic requirement for a control system with infinite state space, and for NCS a key issue is how much information does a feedback controller need in order to stabilize the system. Questions of this kind have motivated much of the study of NCS: stability under communication constraints of linear control systems is studied by Wong and Brockett [105,106], Tatikonda and Mitter [97,98], Elia and Mitter [34], Nair and Evans [74], Liberzon [63] and many others; stability of nonlinear control systems is further studied in [75] and [64].

Stability considerations are important in the analysis of optimal control over the infinite horizon. The study of reachability and stabilization of switched linear systems has attracted considerable interest recently [39,65,95,96,107,108]. Necessary and sufficient conditions for stabilizability for the continuous-time counterpart of (4.18) are obtained in [96] and [109]. Switched discrete-time linear systems are studied in [39], [108] and [110] under different scenarios. We start with the following definition.

Definition 4.4.1. *The stochastic system* (4.1)–(4.2) *is* uniformly stabilizable*, if there exist an admissible policy $v \in \mathcal{V}$, such that*

$$\mathbb{E}^{v}_{X_0}[X_t] \xrightarrow[t\to\infty]{} 0\,, \quad \textit{and} \quad \sup_{t\geq 0}\, \mathbb{E}^{v}_{X_0}\|X_t\|^2 < \infty\,,$$

for any initial condition $X_0 \sim \mathcal{N}(\bar{x}_0, \Sigma_0)$. A policy v having this property is called stable*.*

We begin by discussing the deterministic system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t+1) &= C_{q(t)} x_{t+1} \, , \end{aligned} \tag{4.18}$$

whose state, observation and controls live in the same Euclidean spaces as (4.1)–(4.2), and the pair $(u(t), q(t))$ is chosen as a function of $\{y(1), \dots, y(t)\}$. Let

$$\bar{C} \triangleq \left[C_1^\mathsf{T} \mid C_2^\mathsf{T} \mid \cdots \mid C_\ell^\mathsf{T} \right]^\mathsf{T} .$$

Then, a necessary condition for the existence of a control $\{u(t), q(t)\}$ such that the closed loop system is asymptotically stable to the origin is that the pair (A, B) be stabilizable and the pair $(\bar{C}, A)$ be detectable. This condition is also sufficient, as shown in the following theorem.

Theorem 4.4.1. *Suppose (A, B) is stabilizable and $(\bar{C}, A)$ is detectable and $K \in \mathbb{R}^{N_u \times N_x}$ is such that the matrix $A - BK$ is stable, i.e., has its eigenvalues in the open unit disc of the complex plane. Then, there exist a collection of matrices $\{L_q \, , \; q \in \boldsymbol{Q}\}$, and a sequence $\{q(0), q(1), \dots\}$ such that the controlled system (4.18), under the dynamic feedback control $u(t) = -K\hat{x}(t)$, with*

$$\hat{x}(t+1) = (A - BK)\hat{x}(t) + L_{q(t)}\big(y(t+1) - C_{q(t)}x(t+1)\big) \, ,$$

is uniformly geometrically stable to the origin.

Proof. It is enough to show that the system

$$\hat{x}(t+1) = (A - L_{q(t)} C_{q(t)})\hat{x}(t)$$

is uniformly geometrically stable to the origin. Consider first the case $\ell = 2$, that lends itself to simpler notation. Without loss of generality assume $(\bar{C}, A)$ is observable. Then, there exist row vectors k_q of dimension N_y such that with $c_q = k_q C_q$, we have

$$\mathbb{R}^{N_x} = \mathcal{R}\big[c_1^\mathsf{T} \mid A^\mathsf{T} c_1^\mathsf{T} \mid \cdots \mid (A^{n_1})^\mathsf{T} c_1^\mathsf{T}\big] \oplus \mathcal{R}\big[c_2^\mathsf{T} \mid A^\mathsf{T} c_2^\mathsf{T} \mid \cdots \mid (A^{n_2})^\mathsf{T} c_2^\mathsf{T}\big] \, , \tag{4.19}$$

and $n_1 + n_2 = N_x + 2$. With respect to the ordered basis in (4.19), A and c_1, c_2 take the form

$$TA^{\mathsf{T}}T^{-1} = \begin{bmatrix} \tilde{A}_1 & \tilde{A}_{12} \\ 0 & \tilde{A}_2 \end{bmatrix}, \quad \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} T^{-1} = \begin{bmatrix} \tilde{c}_1 & 0 \\ 0 & \tilde{c}_2 \end{bmatrix},$$

with $\tilde{A}_{12} \in \mathbb{R}^{n_1 \times n_2}$, $\tilde{A}_q \in \mathbb{R}^{n_q \times n_q}$, $\tilde{c}_q^{\mathsf{T}} \in \mathbb{R}^{n_q}$, and the pair $(\tilde{c}_q, \tilde{A}_q)$ is observable, for $q \in \{1, 2\}$.

Let $\gamma > 0$ be such that

$$\max_{q \in \{1,2\}} \{|\lambda| : \lambda \in \sigma(\tilde{A}_q)\} < \gamma\,, \tag{4.20}$$

where σ denotes the spectrum of the matrix, and let $\tilde{\gamma} > 0$ be defined by

$$\tilde{\gamma} \triangleq \tfrac{1}{2} \min\{\gamma, \gamma^{-1}\}\,. \tag{4.21}$$

Select gains $l_q \in \mathbb{R}^{n_q}$, such that

$$\max_{q \in \{1,2\}} \{|\lambda| : \lambda \in \sigma(\tilde{A}_q + l_q \tilde{c}_q)\} < \tilde{\gamma}\,. \tag{4.22}$$

Then, if we let

$$L_1 \triangleq T^{-1} \begin{bmatrix} l_1 \\ 0 \end{bmatrix} k_1\,, \qquad L_2 \triangleq T^{-1} \begin{bmatrix} 0 \\ l_2 \end{bmatrix} k_2\,,$$

and $\bar{A}_q \triangleq \tilde{A}_q + l_q \tilde{c}_q$, we obtain

$$\hat{A}_1 \triangleq T(A + L_1 C_1)T^{-1} = \begin{bmatrix} \bar{A}_1 & \tilde{A}_{12} \\ 0 & \tilde{A}_2 \end{bmatrix},$$

$$\hat{A}_2 \triangleq T(A + L_2 C_2)T^{-1} = \begin{bmatrix} \tilde{A}_1 & \tilde{A}_{12} \\ 0 & \bar{A}_2 \end{bmatrix}.$$

Expressing $x \in \mathbb{R}^n$ in block form $x = \binom{x_1}{x_2}$, with $x_q \in \mathbb{R}^{n_q}$, $q \in \{1, 2\}$, we define the block norm

$$\|x\|_* = \max_{q \in \{1,2\}} \{\|x_q\|\}\,.$$

By (4.20) and (4.22), there exists $M>0$ such that for all $k\in\mathbb{N}$,

$$\|\tilde{A}_q^k\| \le M\gamma^k\,, \quad \|\bar{A}_q^k\| \le M\tilde{\gamma}^k\,, \quad q\in\{1,2\}\,, \tag{4.23}$$

and $\|\tilde{A}_{12}\|\le M$. For $k\in\mathbb{N}$ and $\bar{x}\in\mathbb{R}^n$, we have

$$\hat{A}_1^k\hat{A}_2^k\bar{x} = \begin{pmatrix} \bar{A}_1^k\tilde{A}_1^k\bar{x}_1 + \mathit{\Gamma}_k\bar{x}_2 \\ \tilde{A}_2^k\bar{A}_2^k\bar{x}_2 \end{pmatrix}\,, \tag{4.24}$$

where

$$\mathit{\Gamma}_k \triangleq \sum_{i=0}^{k-1}\left(\bar{A}_1^k\tilde{A}_1^{k-1-i}\tilde{A}_{12}\bar{A}_2^i + \bar{A}_1^{k-1-i}\tilde{A}_{12}\tilde{A}_2^i\bar{A}_2^k\right).$$

Using (4.23) and (4.21), we calculate the following estimates

$$\begin{aligned} \|\bar{A}_1^k\tilde{A}_1^k\bar{x}_1\| &\le M^2\tilde{\gamma}^k\gamma^k\|\bar{x}_1\| \le M^2\left(\tfrac{1}{2}\right)^k\|\bar{x}_1\| \\ \|\tilde{A}_2^k\bar{A}_2^k\bar{x}_2\| &\le M^2\tilde{\gamma}^k\gamma^k\|\bar{x}_2\| \le M^2\left(\tfrac{1}{2}\right)^k\|\bar{x}_2\|\,, \end{aligned} \tag{4.25}$$

and

$$\begin{aligned} \Big\|\bar{A}_1^k\sum_{i=0}^{k-1}\tilde{A}_1^{k-1-i}\tilde{A}_{12}\bar{A}_2^i\bar{x}_2\Big\| &\le M\tilde{\gamma}^k\sum_{i=0}^{k-1}M^3\gamma^{k-1-i}\tilde{\gamma}^i\|\bar{x}_2\| \\ &\le 2M^4\gamma^{-1}\tilde{\gamma}^k\gamma^k\|\bar{x}_2\| \\ &\le 2M^4\gamma^{-1}\left(\tfrac{1}{2}\right)^k\|\bar{x}_2\|\,, \end{aligned} \tag{4.26}$$

$$\begin{aligned} \Big\|\sum_{i=0}^{k-1}\bar{A}_1^{k-1-i}\tilde{A}_{12}\tilde{A}_2^i\bar{A}_2^k\bar{x}_2\Big\| &\le \sum_{i=0}^{k-1}M^4\tilde{\gamma}^{k-1-i}\gamma^i\tilde{\gamma}^k\|\bar{x}_2\| \\ &= \sum_{i=0}^{k-1}M^4\tilde{\gamma}^i\gamma^{k-1-i}\tilde{\gamma}^k\|\bar{x}_2\| \\ &\le 2M^4\gamma^{-1}\left(\tfrac{1}{2}\right)^k\|\bar{x}_2\|\,. \end{aligned} \tag{4.27}$$

It follows by (4.24)–(4.27), that if we select

$$\hat{k} > \log_2\left(2M^2 + 8M^4\gamma^{-1}\right)\,,$$

then

$$\|\hat{A}_1^{\hat{k}}\hat{A}_2^{\hat{k}}\bar{x}\|_* < \tfrac{1}{2}\,\|\bar{x}\|_*\,, \quad \forall \bar{x} \in \mathbb{R}^n\,.$$

Therefore, the periodic switching

$$q(t) = \begin{cases} 1 & t \in \{(2l+1)\hat{k}, \dots, (2l+2)\hat{k}-1\} \\ 2 & t \in \{2l\hat{k}, \dots, (2l+1)\hat{k}-1\} \end{cases}\,,$$

for $l = 0, 1, \dots,$ yields an asymptotically stable system. The general case $\ell \geq 2$, follows in exact analogy: one shows that the map $\hat{A}_1^{\hat{k}}\hat{A}_2^{\hat{k}} \cdots \hat{A}_\ell^{\hat{k}}$ is a contraction with respect to the block norm $\|\cdot\|_*$, for some $\hat{k} \in \mathbb{N}$. Thus, there exists a periodic switching sequence which is stabilizing. □

Theorem 4.4.1 can be applied to characterize the uniform stabilization of (4.1)–(4.2). Consider dynamic output feedback of the form:

$$\begin{aligned} Z_{t+1} &= (A - L_{Q_{t-1}}C_{Q_{t-1}})Z_t + L_{Q_{t-1}}Y_t\,, \quad Z_0 = 0 \\ U_t &= -KZ_t\,, \end{aligned} \tag{4.28}$$

and let $\tilde{Z} \triangleq X_t - Z_t$. Then, by (4.1)–(4.2) and (4.28), we obtain

$$\begin{aligned} X_{t+1} &= (A - BK)X_t - BK\tilde{Z}_t + DW_t \\ \tilde{Z}_{t+1} &= (A - L_{Q_{t-1}}C_{Q_{t-1}})\tilde{Z}_t + (D - L_{Q_{t-1}}F)W_t\,. \end{aligned}$$

By Theorem 4.4.1, provided (A, B) is observable and $(\bar{C}, A)$ is stabilizable, there exist gain matrices K, and $\{L_q\,,\ q \in \boldsymbol{Q}\}$, and a periodic sequence $\{Q_t\}$, such that under this policy (i.e., with $U_t = -KZ_t$), which is denoted by $v_s \in \mathcal{V}$, we have $\mathbb{E}^{v_s}[X_t] \xrightarrow[t\to\infty]{} 0$ and $\mathbb{E}^{v_s}\|X_t\|^2$ is bounded. Furthermore, since by the proof of Theorem 4.4.1, the product $\prod_{t=0}^{n}(A - L_{Q_t}C_{Q_t})$ decays geometrically in norm, there exist constants $\gamma_s \in (0, 1)$ and $M_s > 0$, such that

$$\begin{aligned} \|\mathbb{E}^{v_s}X_t\|^2 &\leq M_s\gamma_s^t\|\bar{x}_0\|^2\,, \qquad \forall t \geq 0 \\ \mathbb{E}^{v_s}\|X_t\|^2 &\leq M_s[\gamma_s^t\,\mathrm{tr}(\varSigma_0) + 1/2]\,, \qquad \forall t \geq 0\,. \end{aligned} \tag{4.29}$$

Note also that under this policy, $\mathbb{E}^{v_s}\|U_t\|^2$ remains bounded, and redefining M_s as the largest of the two bounds, in addition to (4.29), we have

$$\mathbb{E}^{v_s}\|U_t\|^2 \le M_s\,, \qquad \forall t \ge 0\,. \tag{4.30}$$

Remark 4.4.1. *The result in Theorem 4.4.1 can be interpreted in another way: as long as the system is uniformly stabilizable with information all the sensors, the system is uniformly stabilizable via sensor querying, namely, accessing only one sensor's data at a time.*

Remark 4.4.2. *In* (4.28) *the control* U_t *is* $\mathcal{Y}^{t-1}$*-adapted, whereas in* (4.1)–(4.2) *admissible controls* v_t *are defined as* $\mathcal{Y}^t$*-adapted. However, there is no discrepancy: on the one hand, sufficiency is not affected, while on the other* (A, B) *observable and* $(\bar{C}, A)$ *stabilizable is necessary for* (4.1)–(4.2) *to be uniformly stabilizable.*

4.5 Optimal control over the infinite horizon

In this section we study the optimal control problem over the infinite horizon. We are particularly interested in the ergodic control problem, and we approach this via the β-discounted one.

4.5.1 The β-discounted cost

Let $\beta \in (0, 1)$ be the discount factor. For a policy $v \in \mathcal{V}$ define

$$J^v_\beta(\bar{x}_0, \Sigma_0) \triangleq \mathbb{E}^v_{X_0}\left[\sum_{t=0}^{\infty} \beta^t \big(c(Q_t) + r(X_t, U_t)\big)\right],$$

and let $J^*_\beta \triangleq \inf_{v\in\mathcal{V}} \; J^v_\beta$.

Provided (A, B) is stabilizable, and $(\tilde{C}, A)$ is detectable, J^*_β is finite. Indeed, since $J^*_\beta \le J^{v_s}_\beta$, with $v_s \in \mathcal{V}$ the policy in (4.29), an easy calculation shows that there exists a constant $\tilde{M}$ such that

$$\begin{aligned} J^*_\beta(\bar{x}_0, \Sigma_0) &\le \tilde{M}\left(\frac{\|x_0\|^2 + \mathrm{tr}(\Sigma_0)}{1-\beta\gamma_s} + (1-\beta)^{-1}\right) \\ &\le \tilde{M}\left(\frac{1}{1-\beta} + \frac{\|x_0\|^2 + \mathrm{tr}(\Sigma_0)}{1-\gamma_s}\right). \end{aligned} \tag{4.31}$$

The existence and characterization of stationary optimal policies for the β-discounted control problem is the topic of the following theorem.

Theorem 4.5.1. *For the control system* (4.1)–(4.2)*, assume that* (A, B) *is stabilizable, and* $(\bar{C}, A)$ *is detectable. Then there exists a unique positive definite solution* $\mathit{\Pi}^*_\beta$ *to the algebraic Riccati equation*

$$\mathit{\Pi}^*_\beta = R + \beta A^\mathsf{T} \mathit{\Pi}^*_\beta A - \beta^2 A^\mathsf{T} \mathit{\Pi}^*_\beta B (S + \beta B^\mathsf{T} \mathit{\Pi}^*_\beta B)^{-1} B^\mathsf{T} \mathit{\Pi}^*_\beta A \,. \tag{4.32}$$

Define the functional map S_β*,* $\beta \in (0, 1]$*, by*

$$S_\beta(f)(\hat{\mathit{\Pi}}) \triangleq \min_q \left\{ c(q) + \operatorname{tr}(\tilde{\mathit{\Pi}}^*_\beta \hat{\mathit{\Pi}}) + \beta f(\mathcal{T}_q(\hat{\mathit{\Pi}})) \right\}, \tag{4.33}$$

where $\tilde{\mathit{\Pi}}^*_\beta \triangleq R - \mathit{\Pi}^*_\beta + \beta A^\mathsf{T} \mathit{\Pi}^*_\beta A$*. Let*

$$U^*_t = -(S + \beta B^\mathsf{T} \mathit{\Pi}^*_\beta B)^{-1} \beta B^\mathsf{T} \mathit{\Pi}^*_\beta A \hat{X}_t \,, \tag{4.34}$$

where

$$\hat{X}_{t+1} = A\hat{X}_t + BU_t + \hat{K}_{q^*(\hat{\mathit{\Pi}}_t)}(\hat{\mathit{\Pi}}_t)\big(Y_{t+1} - C_{q^*(\hat{\mathit{\Pi}}_t)}(A\hat{X}_t + BU_t)\big) \,, \tag{4.35}$$

with

$$\hat{K}_q(\hat{\mathit{\Pi}}) \triangleq \mathit{\Xi}(\hat{\mathit{\Pi}}) C_q^\mathsf{T} \big(C_q \mathit{\Xi}(\hat{\mathit{\Pi}}) C_q^\mathsf{T} + FF^\mathsf{T} \big)^{-1} \tag{4.36a}$$

$$\hat{\mathit{\Pi}}_{t+1} = \mathcal{T}_{q^*(\hat{\mathit{\Pi}}_t)}(\hat{\mathit{\Pi}}_t) \,, \qquad \hat{\mathit{\Pi}}_0 = \mathit{\Sigma}_0 \,. \tag{4.36b}$$

Let

$$\tilde{J}^{U^*}_\beta(\bar{x}_0, \mathit{\Sigma}_0) \triangleq \bar{x}_0^\mathsf{T} \mathit{\Pi}^*_\beta \bar{x}_0 + \operatorname{tr}(\mathit{\Pi}^*_\beta \mathit{\Sigma}_0) + \beta(1-\beta)^{-1} \operatorname{tr}(\mathit{\Pi}^*_\beta D D^\mathsf{T}) \,. \tag{4.37}$$

There exists a lower semicontinuous $f^*_\beta : \mathcal{M}^+_0 \to \mathbb{R}_+$ *satisfying*

$$f^*_\beta(\hat{\mathit{\Pi}}) = S_\beta(f^*_\beta)(\hat{\mathit{\Pi}}) \,, \tag{4.38}$$

such that if $q^*_\beta : \mathcal{M}^+_0 \to \boldsymbol{Q}$ *is a selector of the minimizer in* (4.33)*, with* $f = f^*_\beta$*, then* $v^*_\beta = (\{U^*_t\}, q^*_\beta)$ *is optimal for the discounted control problem, and for each*

$\beta \in (0,1)$, *the optimal discounted cost is given by*

$$J^*_\beta(\bar{x}_0, \varSigma_0) = \tilde{J}^{U^*}_\beta(\bar{x}_0, \varSigma_0) + f^*_\beta(\varSigma_0)\,. \tag{4.39}$$

Proof. It is well known that, provided (A, B) is stabilizable, the matrix recursive iteration (4.5) for $\varPi_t$ converges to a positive definite matrix $\varPi^*_\beta$ satisfying (4.32). Moreover, (4.32) has a unique solution in $\mathcal{M}^+$. Consider the finite-horizon problem with initial condition $X_0 \sim \mathcal{N}(\bar{x}_0, \varSigma_0)$

$$J^v_k(\bar{x}_0, \varSigma_0) \triangleq \mathbb{E}^v_{X_0}\left[\sum_{t=0}^{k-1} \beta^t \big(c(Q_t) + r(X_t, U_t)\big) + \beta^k X_k^\mathsf{T} \varPi^*_\beta X_k\right], \quad k \in \mathbb{N}\,.$$

It follows by Section 4.3 that the optimal cost J^*_k is given by

$$J^*_k(\bar{x}_0, \varSigma_0) = \bar{x}_0^\mathsf{T} \varPi^*_\beta \bar{x}_0 + \operatorname{tr}(\tilde{\varPi}^*_\beta \varSigma_0) + \sum_{t=1}^{k} \beta^t \operatorname{tr}(\varPi^*_\beta D D^\mathsf{T}) + f_0^{(k)}(\varSigma_0)\,, \tag{4.40}$$

where $f_0^{(k)} : \mathcal{M}^+_0 \to \mathbb{R}$ satisfies

$$f_0^{(k+1)}(\hat{\varPi}) = \min_q \left\{c(q) + \operatorname{tr}(\tilde{\varPi}^*_\beta \hat{\varPi}) + \beta f_0^{(k)}(\mathcal{T}_q(\hat{\varPi}))\right\}, \tag{4.41}$$

with $f_0^{(0)} = 0$. Since $J^*_k \le J^{v_s}_k$, where v_s is the policy in Theorem 4.4.1, it follows that $\{f_0^{(k)}\}$ is bounded pointwise in $\mathcal{M}^+_0$. Since, in addition, $f_0^{(k)} \uparrow$, it converges to a lower semicontinuous function f^*_β, and taking monotone limits, (4.41) yields (4.38). Since f^*_β is locally bounded, it follows that $\beta^t \mathbb{E}^{q^*_\beta}_{X_0}[\operatorname{tr}(\hat{\varPi}_t)] \to 0$ as $t \to \infty$. Thus, the estimate in (4.31) yields $\beta^t \mathbb{E}^{q^*_\beta}_{X_0}[f^*_\beta(\hat{\varPi}_t)] \to 0$, as $t \to \infty$. Using the dynamic programming equation (4.38), we have

$$f^*_\beta(\varSigma_0) = \mathbb{E}^{q^*_\beta}_{X_0}\left[\sum_{k=0}^{t-1} \beta^k \big(c(Q_k) + \operatorname{tr}(\tilde{\varPi}^*_\beta \hat{\varPi}_k)\big)\right] + \beta^t \mathbb{E}^{q^*_\beta}_{X_0}[f^*_\beta(\hat{\varPi}_t)]\,,$$

and taking limits as $t \to \infty$, we obtain

$$f^*_\beta(\varSigma_0) = \mathbb{E}^{q^*_\beta}_{X_0}\left[\sum_{k=0}^{\infty} \beta^k \big(c(Q_k) + \operatorname{tr}(\tilde{\varPi}^*_\beta \hat{\varPi}_k)\big)\right].$$

One more application of (4.38) shows that for all $v \in \mathcal{V}$, such that $J^v_\beta(\bar{x}_0, \Sigma_0) < \infty$,

$$f^*_\beta(\Sigma_0) \le \mathbb{E}^v_{X_0}\left[\sum_{k=0}^{\infty} \beta^k \bigl(c(Q_k) + \operatorname{tr}(\tilde{\Pi}^*_\beta \hat{\Pi}_k)\bigr)\right],$$

and thus, q^*_β is optimal. The proof is complete. $\square$

4.5.2 The average cost

Before proceeding to the analysis of the ergodic control problem, we establish some useful properties of the Riccati map $\mathcal{T}_q$ defined in (4.9). We have the identity

$$\mathcal{T}_q(\hat{\Pi}) = (I - \hat{K}_q(\hat{\Pi})C_q)\Xi(\hat{\Pi})(I - \hat{K}_q(\hat{\Pi})C_q)^\mathsf{T} + \hat{K}_q(\hat{\Pi})FF^\mathsf{T}\hat{K}^\mathsf{T}_q(\hat{\Pi})\,. \tag{4.42}$$

For $\varepsilon > 0$, let

$$\mathcal{M}^+_\varepsilon \triangleq \{\hat{\Pi} \in \mathcal{M}^+ : \min\{\lambda \mid \lambda \in \sigma(\hat{\Pi})\} > \varepsilon\}\,.$$

We define the operator $\circ$: $(f \circ g)(x) \triangleq f(g(x))$. To prove the existence of a stationary optimal policy for the ergodic control problem, we employ Lemmas 4.5.2–4.5.3 below.

Lemma 4.5.2. *There exists $\varepsilon > 0$ and $\kappa \in \mathbb{N}$, such that*

$$\mathcal{T}_{q_{\kappa-1}} \circ \mathcal{T}_{q_{\kappa-2}} \circ \cdots \circ \mathcal{T}_{q_0}(0) \in \mathcal{M}^+_\varepsilon\,,$$

for every sequence $\{q_0, q_1, \dots, q_{\kappa-1}\}$, if and only if the pair (A, D) is controllable, in which case $\kappa \le N_x$.

Proof. Suppose (A, D) is not controllable. Then there exists an eigenvector $z \in \mathbb{R}^{N_x}$ of A^T such that $z^\mathsf{T} D = 0$. By (4.36a) and (4.42), $z^\mathsf{T}\mathcal{T}_q(0) = z^\mathsf{T} DD^\mathsf{T} = 0$. Proceeding by induction, suppose that for any sequence $\{q_0, q_1, \dots, q_{k-1}\}$ of length $k \ge 1$,

$$z^\mathsf{T}\mathcal{T}_{q_{k-1}} \circ \cdots \circ \mathcal{T}_{q_0}(0) = 0\,. \tag{4.43}$$

Set $\hat{\Pi}_k = \mathcal{T}_{q_{k-1}} \circ \cdots \circ \mathcal{T}_{q_0}(0)$. Then, since z is an eigenvector of A^T, $z^\mathsf{T} A\hat{\Pi}_k = 0$, and it follows by (4.36a) that $z^\mathsf{T}\hat{K}_q(\hat{\Pi}_k) = 0$, for all $q \in \boldsymbol{Q}$. In turn, by (4.42), $z^\mathsf{T}\mathcal{T}_q(\hat{\Pi}_k) = 0$, for all $q \in \boldsymbol{Q}$, and the induction is complete.

It is straightforward to verify that $\hat{\Pi}' \geq \hat{\Pi}$ implies $\mathcal{T}_q(\hat{\Pi}') \geq \mathcal{T}_q(\hat{\Pi})$ (see [53]). Hence, to show the converse, suppose $\hat{\Pi}_k = \mathcal{T}_{q_{k-1}} \circ \cdots \circ \mathcal{T}_{q_0}(0)$ is singular for all sequences $\{q_0, q_1, \ldots, q_{k-1}\}$ of length $k \leq N_x$. Let $\ker(\hat{\Pi}_k)$ denote the null-space of $\hat{\Pi}_k$, and denote by z_k its arbitrary element. As already mentioned, $\ker(\hat{\Pi}_k) \subset \ker(\hat{\Pi}_{k-1})$, for all $k \geq 1$. Therefore, using also the assumption that FF^T is nonsingular, successive applications of (4.42) yield

$$z_k^\mathsf{T} \hat{K}_{q_j}(\hat{\Pi}_j) = 0\,, \quad z_k^\mathsf{T} D = 0\,, \quad \text{for all } k > j \geq 0\,, \tag{4.44}$$

and $z_k^\mathsf{T} A \hat{\Pi}_{k-1} = 0$, which implies

$$A^\mathsf{T} \ker(\hat{\Pi}_k) \subset \ker(\hat{\Pi}_{k+1})\,, \quad \forall k \geq 1\,. \tag{4.45}$$

Using (4.36a), (4.42) and (4.44)–(4.45), we can show that

$$\ker(\hat{\Pi}_k) = (A^\mathsf{T})^{-1} \ker(\hat{\Pi}_{k-1}) \bigcap \ker(DD^\mathsf{T})\,. \tag{4.46}$$

for all $k \geq 1$. It follows from (4.46), that $\ker(\hat{\Pi}_j) = \ker(\hat{\Pi}_{N_x})$, for all $j \geq N_x$, and that if $\hat{\Pi}_{N_x}$ is singular, then (A, D) is not controllable. $\square$

Lemma 4.5.3. *The functions $\mathcal{T}_q : \mathcal{M}_0^+ \to \mathcal{M}_0^+$ and $f_\beta^* : \mathcal{M}_0^+ \to \mathbb{R}_+$ are concave.*

Proof. Note that if the filtering at time t is based upon y^{t-1} instead of y^t, the corresponding Riccati map is different from $\mathcal{T}$ and its convexity has been shown in [93].

To prove the concavity of $\mathcal{T}_q$, we show that for any scalar θ and symmetric square matrix Z, $\frac{\partial^2 \mathcal{T}_q(\hat{\Pi}+\theta Z)}{\partial \theta^2} \leq 0$. To simplify the notation, we define

$$\hat{\Pi}' \triangleq \hat{\Pi} + \theta Z\,, \qquad \Lambda_q \triangleq (C_q \Xi(\hat{\Pi}') C_q^\mathsf{T} + FF^\mathsf{T})^{-1}\,.$$

After some algebra, we obtain

$$\frac{\partial^2 \mathcal{T}_q(\hat{\Pi}')}{\partial \theta^2} = -2(C_q^\mathsf{T} \Lambda_q C_q \Xi - I)^\mathsf{T} A Z^\mathsf{T} A^\mathsf{T} C_q^\mathsf{T} \Lambda_q C_q A Z A^\mathsf{T} (C_q^\mathsf{T} \Lambda_q C_q \Xi - I)\,.$$

Since $\Lambda_q > 0$ and Z is symmetric, we have $\frac{\partial^2 \mathcal{T}_q(\hat{\Pi}')}{\partial \theta^2} \le 0$, which shows that $\mathcal{T}_q$ is concave. The concavity of f^*_β follows from the fact that the map S_β is concavity-conserving, namely, $S_\beta(f)$ is concave if f is concave. $\square$

To characterize the ergodic control problem, we adopt the vanishing discount method, i.e., an asymptotic analysis as the discount factor $\beta \to 1$. By (4.37)–(4.39), for any $\bar{x}_1$, $\bar{x}_2$ in $\mathbb{R}^{N_x}$ and Σ_1, Σ_2 in $\mathcal{M}_0^+$,

$$f^*_\beta(\Sigma_1) - f^*_\beta(\Sigma_2) = J^*_\beta(\bar{x}_1, \Sigma_1) - J^*_\beta(\bar{x}_2, \Sigma_2) - (\bar{x}_1^\mathsf{T} \Pi^*_\beta \bar{x}_1 - \bar{x}_2^\mathsf{T} \Pi^*_\beta \bar{x}_2) - \operatorname{tr}\bigl(\Pi^*_\beta(\Sigma_1 - \Sigma_2)\bigr)\,. \tag{4.47}$$

Also using

$$J^*_\beta(\bar{x}_0, \Sigma_0) \le \mathbb{E}^{v_s}_{X_0}\left[\sum_{t=0}^{k-1} \beta^t \bigl(c(Q_t) + r(X_t, U_t)\bigr)\right] + \beta^k \mathbb{E}^{v_s}_{X_0}[J^*_\beta(\hat{X}_k, \hat{\Pi}_k)]\,,$$

we obtain, that for some constant $\tilde{M}' > 0$, and M_s the constant in (4.29),

$$J^*_\beta(\bar{x}_0, \Sigma_0) \le \tilde{M}'\bigl(1 + \|x_0\|^2 + \operatorname{tr}(\Sigma_0)\bigr) + \sup_{\substack{\|x\|^2 \le M_s \\ \operatorname{tr}(\Sigma) \le M_s}} J^*_\beta(\bar{x}, \Sigma)\,. \tag{4.48}$$

Let $\mathcal{B}_s \subset \mathbb{R}^{N_x} \times \mathcal{M}_0^+$ be a bounded ball in $\mathcal{M}_0^+$ containing the set $\{\Sigma : \operatorname{tr}(\Sigma) \le M_s\}$, and such that

$$\mathcal{T}_{q_\kappa} \circ \mathcal{T}_{q_{\kappa-1}} \circ \cdots \circ \mathcal{T}_{q_0}(0) \in \mathcal{B}_s\,, \quad \forall\, \{q_0, q_1, \dots, q_\kappa\} \subset \boldsymbol{Q}^{\kappa+1}\,.$$

Since f^*_β depends only on Σ, and since Π^*_β converges to a limit in $\mathcal{M}^+$, as $\beta \to 1$, it follows from (4.47) and (4.48) that there exists a continuous function $G_s : \mathbb{R}_+ \to \mathbb{R}_+$, having affine growth, such that

$$f^*_\beta(\Sigma) - \sup_{\Sigma' \in \mathcal{B}_s} \{f^*_\beta(\Sigma')\} \le G_s(\operatorname{tr}(\Sigma))\,, \tag{4.49}$$

for all $\Sigma \in \mathcal{M}_0^+$. Define

$$\bar{f}_\beta \triangleq f^*_\beta(\Sigma) - \inf_{\Sigma' \in \mathcal{M}_0^+} f^*_\beta(\Sigma')$$

$$\operatorname*{span}_{\mathcal{B}_s}(f^*_\beta) \triangleq \sup_{\Sigma \in \mathcal{B}_s} f^*_\beta(\Sigma) - \inf_{\Sigma \in \mathcal{B}_s} f^*_\beta(\Sigma)\,.$$

Equicontinuity of the differential discounted value function $\bar{f}_\beta$, is established in the following lemma.

Lemma 4.5.4. *Under the assumptions of Theorem 4.5.1,*

(i) $\inf_{\Sigma \in \mathcal{M}_0^+} f^*_\beta(\Sigma) = f^*_\beta(0)$, *for any* $\beta \leq 1$.

(ii) Suppose (A, D) *is a controllable pair. Then,* $\bar{f}_\beta$ *is locally bounded, uniformly in* $\beta \in (0, 1)$.

(iii) Provided (A, D) *is a controllable pair,* $\{\bar{f}_\beta\,,\ 0 < \beta < 1\}$ *is equicontinuous on compact subsets of* $\mathcal{M}_0^+$.

Proof. (i) As mentioned in the proof of Lemma 4.5.2, $\Sigma' \geq \Sigma$ implies $\mathcal{T}_q(\Sigma') \geq \mathcal{T}_q(\Sigma)$. Hence, it follows from (4.33) that if $\Sigma' \geq \Sigma$, then $f^*_\beta(\Sigma') \geq f^*_\beta(\Sigma)$. Thus $f^*_\beta(0) = \inf_{\Sigma \in \mathcal{M}_0^+} f^*_\beta(\Sigma)$.

(ii) Let $\varepsilon > 0$ be the constant in Lemma 4.5.2. For a $\beta \in (0, 1)$, let $\Sigma^*_\beta \in \mathcal{B}_s$ be such that

$$f^*_\beta(\Sigma^*_\beta) \geq \sup_{\Sigma \in \mathcal{B}_s} f^*_\beta(\Sigma) - \varepsilon\,.$$

If v^*_β is an optimal β-discounted policy, then

$$\begin{aligned} f^*_\beta(0) &= \mathbb{E}_0^{v^*_\beta}\left[\sum_{t=0}^{\kappa-1} \beta^t (c(Q_t) + \operatorname{tr}(\tilde{\Pi}^*_\beta \hat{\Pi}_t)) + \beta^\kappa f^*_\beta(\hat{\Pi}_\kappa)\right] \\ &\geq \beta^\kappa \mathbb{E}_0^{v^*_\beta}\left[f^*_\beta(\hat{\Pi}_\kappa)\right]. \end{aligned}$$

Thus,

$$\operatorname*{span}_{\mathcal{B}_s}(f^*_\beta) \leq f^*_\beta(\Sigma^*_\beta) - f^*_\beta(0) + \varepsilon$$

$$
\begin{aligned}
&\leq f_\beta^*(\Sigma_\beta^*) - \beta^\kappa \mathbb{E}_0^{\pi_\beta}\big[f_\beta^*(\hat{\Pi}_\kappa)\big] + \varepsilon \\
&= (1-\beta^\kappa) f_\beta^*(\Sigma_\beta^*) + \beta^\kappa \mathbb{E}_0^{\pi_\beta}\big[f_\beta^*(\Sigma_\beta^*) - f_\beta^*(\hat{\Pi}_\kappa)\big] + \varepsilon \\
&\leq (1-\beta^\kappa) f_\beta^*(\Sigma_\beta^*) + \beta^\kappa (1-\varepsilon) \operatorname*{span}_{\mathcal{B}_s}(f_\beta^*) + \varepsilon\,,
\end{aligned}
$$

where the last inequality follows from the concavity of f_β^*, and the fact that $\hat{\Pi}_\kappa \in \mathcal{M}_\varepsilon^+$. Therefore,

$$
\begin{aligned}
\operatorname*{span}_{\mathcal{B}_s}(f_\beta^*) &\leq \frac{(1-\beta^\kappa) f_\beta^*(\Sigma_\beta^*) + \varepsilon}{1-\beta^\kappa(1-\varepsilon)} \\
&= \frac{(1+\beta+\dots+\beta^{\kappa-1})(1-\beta) f_\beta^*(\Sigma_\beta^*) + \varepsilon}{1-\beta^\kappa(1-\varepsilon)} \\
&\leq \frac{\kappa}{\varepsilon}(1-\beta) f_\beta^*(\Sigma_\beta^*) + 1\,.
\end{aligned}
$$

Since, by (4.31) $(1-\beta) f_\beta^*(\Sigma_\beta^*)$ is bounded, uniformly in $\beta \in (0,1)$, the same holds for $\operatorname{span}_{\mathcal{B}_s}(f_\beta^*)$. The result then follows by (4.49).

(iii) Equicontinuity of $\{\bar{f}_\beta\}$ on bounded subsets of $\mathcal{M}_\varepsilon^+$, for any $\varepsilon > 0$, follows from the uniform boundedness and concavity of $\{\bar{f}_\beta\}$ [83]. Since, by (1), $\mathcal{T}_q(\Sigma) \geq \mathcal{T}_q(0)$, for any $\Sigma \in \mathcal{M}_0^+$, then by Lemma 4.5.2,

$$
\mathcal{T}_{q_\kappa} \circ \mathcal{T}_{q_{\kappa-1}} \circ \dots \circ \mathcal{T}_{q_0}(\Sigma) \in \mathcal{M}_\varepsilon^+\,, \quad \forall \Sigma \in \mathcal{M}_0^+\,,
$$

for all $\{q_0, \dots, q_\kappa\} \in \boldsymbol{Q}^{\kappa+1}$. Fix the initial condition Σ, and let $\{q_0, q_1, \dots\}$ be a corresponding β-discounted optimal sequence of queries, i.e., selectors from the minimizer in (4.33). Define $\mathcal{T}_{q_k,\dots,q_0} \triangleq \mathcal{T}_{q_k} \circ \mathcal{T}_{q_{k-1}} \circ \dots \circ \mathcal{T}_{q_0}$. Using (4.38), for any $\Sigma' \in \mathcal{M}_0^+$,

$$
\begin{aligned}
f_\beta^*(\Sigma') - f_\beta^*(\Sigma) \leq \operatorname{tr}(\tilde{\Pi}_\beta^*(\Sigma'-\Sigma)) + \sum_{k=0}^{\kappa-1} \beta^{k+1} \operatorname{tr}\big(\tilde{\Pi}_\beta^*(\mathcal{T}_{q_k,\dots,q_0}(\Sigma') - \mathcal{T}_{q_k,\dots,q_0}(\Sigma))\big) \\
+ \beta^{\kappa+1}\big[\bar{f}_\beta(\mathcal{T}_{q_\kappa,\dots,q_0}(\Sigma')) - \bar{f}_\beta(\mathcal{T}_{q_\kappa,\dots,q_0}(\Sigma))\big]\,. \qquad (4.50)
\end{aligned}
$$

Thus, equicontinuity on every compact subset of $\mathcal{M}_0^+$ follows from (4.50), by exploiting the continuity of $\mathcal{T}_{q_\kappa,\dots,q_0}$, the property $\mathcal{T}_{q_\kappa,\dots,q_0}(\mathcal{M}_0^+) \subset \mathcal{M}_\varepsilon^+$, and the fact that and $\bar{f}_\beta$ is equicontinuous on bounded subsets of $\mathcal{M}_\varepsilon^+$.

□

We now turn to the ergodic control problem. For a policy $v \in \mathcal{V}$ define

$$J^v(\bar{x}_0, \Sigma_0) \triangleq \limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \bigl(c(Q_t) + X_t^{\mathsf{T}} R X_t + U_t^{\mathsf{T}} S U_t\bigr)\,,$$

and let $J^* = \inf_{v\in\mathcal{V}} J^v$. The main result of this section is the following.

Theorem 4.5.5 (Ergodic control)**.** *Assume that (A, B) is stabilizable, $(\bar{C}, A)$ is detectable, and (A, D) is controllable. Define the functional map S, by*

$$S(h)(\hat{\Pi}) \triangleq \min_q \bigl\{c(q) + \operatorname{tr}(\tilde{\Pi}^* \hat{\Pi}) + h(\mathcal{T}_q(\hat{\Pi}))\bigr\}\,, \tag{4.51}$$

where $\tilde{\Pi}^* \triangleq R - \Pi^* + A^{\mathsf{T}}\Pi^* A$, and $\Pi^* \in \mathcal{M}^+$ solves

$$\Pi^* = A^{\mathsf{T}}\Pi^* A + R \\ - A^{\mathsf{T}}\Pi^* B(S + B^{\mathsf{T}}\Pi^* B)^{-1} B^{\mathsf{T}}\Pi^* A\,. \tag{4.52}$$

There exists a nonnegative constant ϱ^* and a continuous $h : \mathcal{M}_0^+ \to \mathbb{R}_+$ satisfying

$$h(\hat{\Pi}) + \varrho^* = S(h)(\hat{\Pi})\,. \tag{4.53}$$

Let $q^* : \mathcal{M}_0^+ \to \boldsymbol{Q}$ be a selector of the minimizer in (4.51)*. Set*

$$U_t^* = -(S + B^{\mathsf{T}}\Pi^* B)^{-1} B^{\mathsf{T}}\Pi^* A\hat{X}_t\,, \tag{4.54}$$

where

$$\hat{X}_{t+1} = A\hat{X}_t + BU_t + \hat{K}_{q^*(\hat{\Pi}_t)}(\hat{\Pi}_t)\bigl(Y_{t+1} \\ - C_{q^*(\hat{\Pi}_t)}(A\hat{X}_t + BU_t)\bigr)\,,$$

with $\hat{K}_q$ as in (4.8a)*, and*

$$\hat{\Pi}_{t+1} = \mathcal{T}_{q^*(\hat{\Pi}_t)}(\hat{\Pi}_t)\,, \qquad \hat{\Pi}_0 = \Sigma_0\,. \tag{4.55}$$

Then, $v^* = (\{U_t^*\}, q^*)$ is optimal for the ergodic control problem, and

$$J^* = \operatorname{tr}(\Pi^* D D^{\mathsf{T}}) + \varrho^* \,.$$

Furthermore, v^* is stable.

Proof. It is well known that, provided (A, B) is stabilizable, Π_β^* converges as $\beta \to 1$ to $\Pi^* \in \mathcal{M}^+$, which is the unique positive definite solution to the algebraic Riccati equation (4.52). Thus it suffices to turn our attention to the query policy. By Lemma 4.5.4, $\{\bar{f}_\beta\}$ is locally equicontinuous and bounded, and thus along some sequence $\beta_k \to 0$, $\bar{f}_{\beta_k}$ converges to some continuous function h, while at the same time $(1-\beta_k) f_{\beta_k}(0)$ converges to some constant ϱ^*. Taking limits in (4.38), we obtain (4.53).

By (4.53), there exists $M_0 > 0$ such that $\operatorname{tr}(\Sigma) > M_0$ implies

$$h(\mathcal{T}_{q^*(\Sigma)}(\Sigma)) - h(\Sigma) < -1 \,, \qquad \forall q \in \boldsymbol{Q} \,.$$

This shows that $\sup_{t \geq 0} \mathbb{E}_{X_0}^{q^*}[\hat{\Pi}_t] < \infty$, for all X_0. Let

$$K^* \triangleq (S + B^{\mathsf{T}} \Pi^* B)^{-1} B^{\mathsf{T}} \Pi^* A \,.$$

Since

$$(A - BK^*)^{\mathsf{T}} \Pi^* (A - BK^*) - \Pi^* = -R - (K^*)^{\mathsf{T}} S K^* \,,$$

it follows that $(A - BK^*)$ is a stable matrix. Thus, from

$$X_{t+1} = (A - BK^*) X_t + BK^*(X_t - \hat{X}_t) + DW_t \,,$$

we obtain

$$\mathbb{E}_{X_0}^{q^*}[X_t] \xrightarrow[t \to \infty]{} 0 \,, \quad \text{and} \quad \sup_{t \geq 0} \mathbb{E}_{X_0}^{q^*} \|X_t\|^2 < \infty \,.$$

By (4.53), for any admissible $\{Q_t\}$,

$$\varrho^* + \frac{h(\Sigma_0) - h(\hat{\Pi}_N)}{N} \leq \frac{1}{N} \sum_{t=0}^{N-1} \bigl(c(Q_t) + \operatorname{tr}(\tilde{\Pi}^* \hat{\Pi}_t)\bigr) \,, \tag{4.56}$$

with equality when $Q_t = q^*$. Since the function G_s in (4.49) has affine growth, it follows that $h(\Sigma) \le G_s^*(\mathrm{tr}(\Sigma))$, for some affine function G_s^*. Therefore, since q^* is stable, $\frac{h(\hat{\Pi}_N)}{N} \to 0$, as $N \to \infty$, which in turn implies by (4.56) that

$$\varrho^* = \limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \bigl(c(Q_t) + \mathrm{tr}(\tilde{\Pi}^* \hat{\Pi}_t)\bigr)\,, \quad \mathbb{P}_{X_0}^{q^*} - \text{a.s.}$$

Also for any policy $v \in \mathcal{V}$ such that the limit supremum of the expectation of the right hand side of (4.56) is finite, we have $\frac{\mathbb{E}_{X_0}^v[h(\hat{\Pi}_{N_k})]}{N_k} \to 0$, along some subsequence $N_k \to \infty$. Thus

$$\liminf_{n\to\infty} \frac{h(\hat{\Pi}_n)}{n} = 0\,, \quad \mathbb{P}_{X_0}^{v} - \text{a.s.} \tag{4.57}$$

By (4.56)–(4.57),

$$\varrho^* \le \limsup_{k\to\infty} \frac{1}{N} \mathbb{E}_{X_0}^v \left[\sum_{t=0}^{N-1} \bigl(c(Q_t) + \mathrm{tr}(\tilde{\Pi}^* \hat{\Pi}_t)\bigr) \right] .$$

Hence q^* is optimal. This completes the proof. □

Remark 4.5.1. *The assumption* (A, D) *controllable cannot be relaxed in general. Lack of this assumption may result in the long-run optimal cost to depend on the initial condition* Σ_0*.*

Remark 4.5.2. *In summary, the steps to compute the optimal controller are as follows: First we solve the Riccati equation* (4.52) *for* $\Pi^* \in \mathcal{M}^+$*. The optimal control is the linear feedback controller in* (4.54) *with a constant gain. Next, we solve the HJB equation* (4.53) *to obtain a stationary optimal policy* q^* *for the query. The optimal query is a function of* $\hat{\Pi}_t$*, and the state estimates are updated according to* (4.55)*.*

4.6 Example: optimal switching estimation

Since the switching control for the observation is the key feature of the problem, the examples presented in this section concentrate on the optimal estimation problem. In other words, the objective is to estimate the system state X_t while minimizing the infinite-horizon criteria with respect to the running cost $\tilde{g}(\hat{\Pi}, q) = c(q) + \mathrm{tr}(\hat{\Pi})$. In this section we present examples of one and two-dimensional systems with a binary query variable, i.e., $\boldsymbol{Q} = \{1, 2\}$.

4.6.1 1-D case

Consider a one-dimensional system as in (4.1)–(4.2), with $C_q \neq 0$, $q \in \{1, 2\}$. If we let $V_q \triangleq F/C_q$, i.e., the normalized noise, $\mathcal{T}_q$ takes the form

$$\mathcal{T}_q(\hat{\Pi}) = A^2\hat{\Pi} + D^2 - \frac{(A^2\hat{\Pi} + D^2)^2}{A^2\hat{\Pi} + D^2 + V_q^2},$$

and the HJB equations for the discounted and ergodic criteria take the form

$$f^*_\beta(\hat{\Pi}) = \min_q \left\{c(q) + \hat{\Pi} + \beta f^*_\beta(\mathcal{T}_q(\hat{\Pi}))\right\} \tag{4.58a}$$

$$\varrho + f^*_\beta(\hat{\Pi}) = \min_q \left\{c(q) + \hat{\Pi} + f^*_\beta(\mathcal{T}_q(\hat{\Pi}))\right\}. \tag{4.58b}$$

Suppose $V_1 > V_2$ and that the cost of observation satisfies $c(1) < c(2)$. In other words, Sensor 1 has a lower sensing capability and lower cost, while Sensor 2 has a higher sensing capability and cost.

Let $\hat{\Pi}^*_1$, $\hat{\Pi}^*_2$ denote the unique fixed points of $\mathcal{T}_1$, $\mathcal{T}_2$, respectively. Since $V_1 > V_2$, we have $\hat{\Pi}^*_1 > \hat{\Pi}^*_2$. The iterates of the map $\mathcal{T}_q$, converge to $\hat{\Pi}^*_q$, hence we restrict our attention to the set of initial conditions $[\hat{\Pi}^*_2, \hat{\Pi}^*_1]$, which is invariant under $\mathcal{T}_q$, $q \in \{1, 2\}$. For $\hat{\Pi} \in (\hat{\Pi}^*_2, \hat{\Pi}^*_1)$, $T_2(\hat{\Pi}) < \hat{\Pi} < T_1(\hat{\Pi})$.

Using the method of successive iterates of the dynamic programming operator, we can derive sharp conditions for the optimal query policy to be switching between the two sensors, and not to be a constant. This is summarized in the following proposition, whose proof is omitted.

Proposition 4.6.1. *Let q^*_β denote the minimizer in* (4.58a)*, for $\beta \in (0, 1)$, and the minimizer in* (4.58b)*, when $\beta = 1$. Then, given $\beta \in (0, 1]$, there exists $\delta > 0$ such that*

(i) $q^*_\beta(\hat{\Pi}) = 1$*, for $\hat{\Pi} \in [\hat{\Pi}^*_2, \hat{\Pi}^*_2 + \delta]$, if and only if*

$$c(2) - c(1) > \sum_{k=0}^{\infty} \beta^k \left(\mathcal{T}_2^k \circ \mathcal{T}_1(\hat{\Pi}^*_2) - \hat{\Pi}^*_2\right).$$

(ii) $q^*_\beta(\hat{\Pi}) = 2$, *for* $\hat{\Pi} \in [\hat{\Pi}^*_1 - \delta, \hat{\Pi}^*_1]$, *if and only if*

$$c(2) - c(1) < \sum_{k=0}^{\infty} \beta^k \left(\hat{\Pi}^*_1 - \mathcal{T}^k_1 \circ \mathcal{T}_2(\hat{\Pi}^*_1)\right).$$

The optimal query policy for the one-dimensional example can be easily obtained numerically by standard algorithms, like value iteration or policy iteration. After running numerous simulations, it appears that the optimal query policy for both the discounted and average costs is a threshold policy, namely, the optimal q^*_β takes the form

$$q^*_\beta = \begin{cases} 1, & \hat{\Pi} < \hat{\Pi}^*_\beta \\ 2, & \hat{\Pi} \geq \hat{\Pi}^*_\beta \end{cases}$$

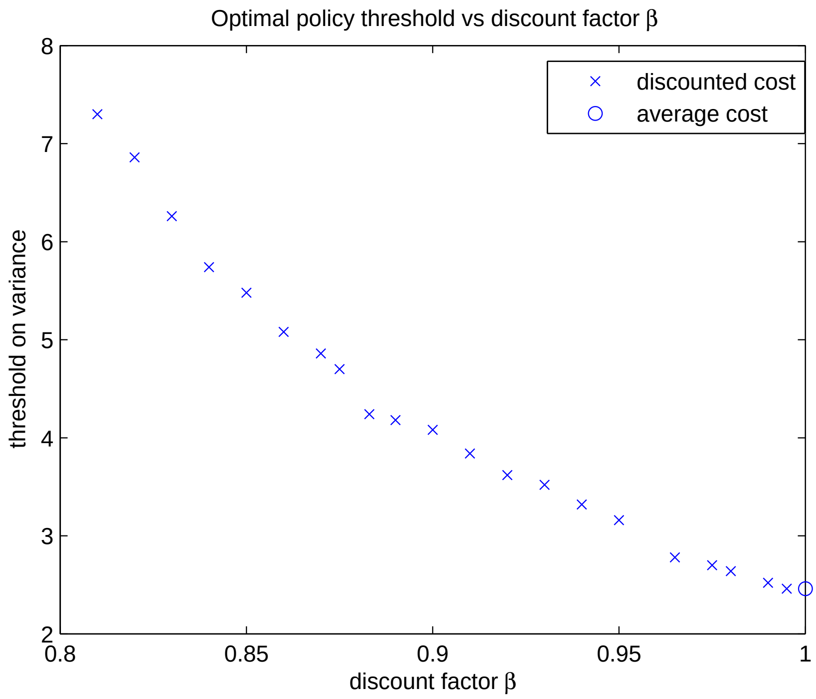


Figure 4.2: The optimal policy threshold vs the discount factor β

The threshold point $\hat{\Pi}^*_\beta$, as a function of the discount factor β, is displayed in Figure 4.2. As β approaches 1, the optimal threshold for the discounted cost converges to that of the average cost. Furthermore, the optimal threshold is a decreasing function of β. This agrees with Proposition 4.6.1, and also agrees with intuition that as the future is weighted more in the criterion, the frequency with which the optimal policy chooses the more accurate and costly observation increases.

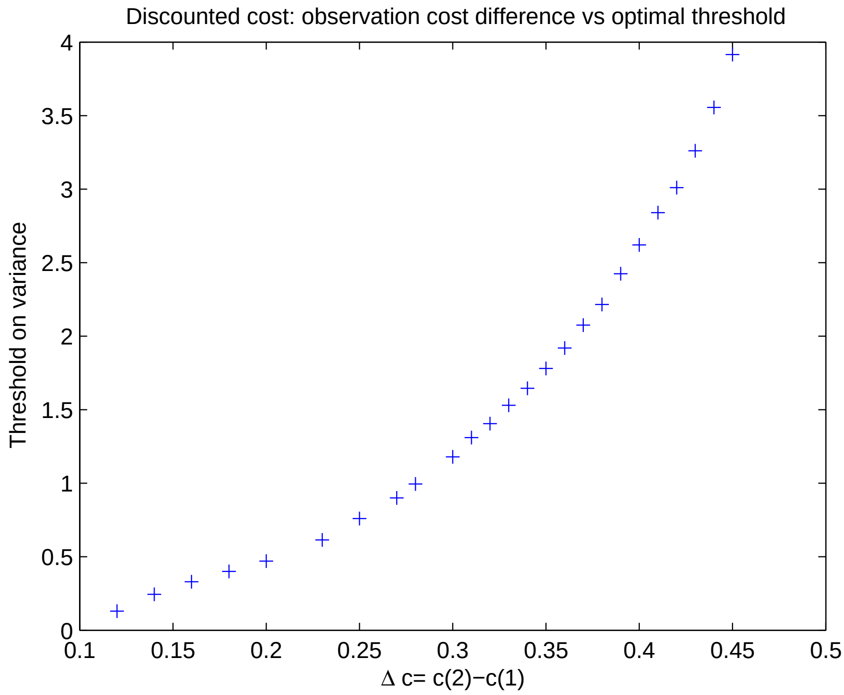


Figure 4.3: The cost difference vs the optimal threshold

Figure 4.3, shows the variation of the optimal threshold as a function of the cost differential. The threshold point is an increasing function of the cost differential and once the latter increases in value beyond 0.45 the optimal policy is a constant, and the controller chooses to use the least costly observation all the time.

4.6.2 2-D case

We present an example of a two-dimensional system with system state $X = [X^1, X^2]^\mathsf{T}$, a scalar observation, and the following parameters:

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \qquad DD^\mathsf{T} = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix}$$
$$F^2 = 0.2 \qquad C_1 = [1 \quad 0]\,, \qquad C_2 = [0 \quad 1]\,.$$

The running cost is $c(1) = c(2) = 0$. Since the pairs (A, C_q) are not detectable, this example can be viewed as a problem of optimal switching estimation.

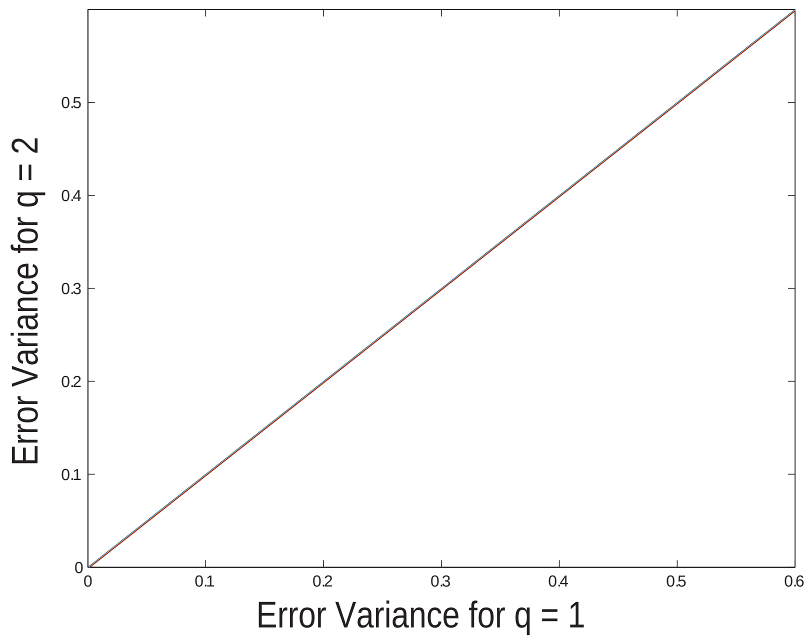


Figure 4.4: The optimal switching curve for the first 2-D example

Figure 4.4 shows the optimal switching curve to minimize the trace of estimation error variance, and can be interpreted as follows: when $\hat{\Pi}^1$, the estimation variance of X^1 is larger than the estimation variance $\hat{\Pi}^2$ of X^2, we query Sensor 1, and vice versa. The switching curve is a straight line due to symmetry.

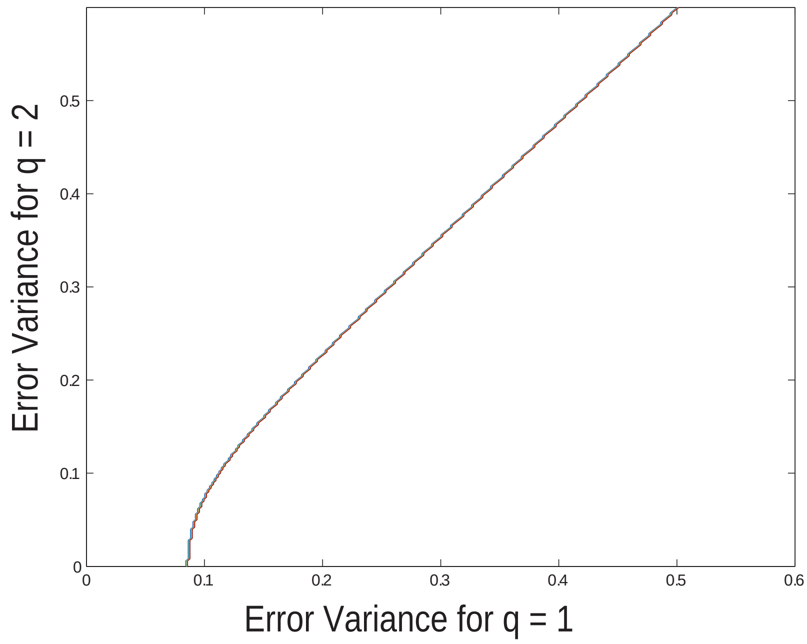


Figure 4.5: The optimal switching curve for the second 2-D example

Next suppose that Sensor 2 has lower observation noise and higher price, i.e.,

$$F_1^2 = 0.1\,, \quad F_2^2 = 0.2\,, \quad c(1) = 0.05\,, \quad c(2) = 0\,,$$

while the rest of the parameters are kept the same as before. This has the following impact on the optimal switching curve, as shown in Figure 4.5: Near the origin, where the penalty on the estimation errors is small, Sensor 2 is used, due to its lower operation cost; far away from the origin, where the estimation error dominates the cost of querying, the symmetry of Figure 4.4 is broken, and Sensor 1 is favored.

In the third 2-D example, both sensors can fully detect the unstable eigenmode of the system state, i.e.,

$$C_1 = [1.0 \;\; 1.0], \qquad C_2 = [1.1 \;\; 1.1]\,,$$

and

$$F_1^2 = 0.2\,, \quad F_2^2 = 0.1\,, \qquad c(1) = 0\,, \quad c(2) = 0.05\,.$$

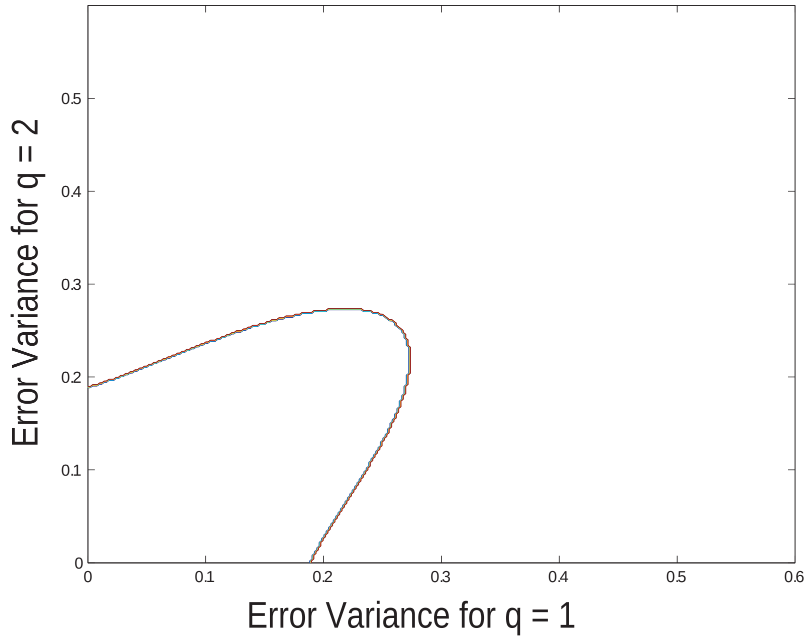


Figure 4.6: The optimal switching curve for the third 2-D example

Figure 4.6 portrays the optimal switching curve for this example. When the estimation error lies in the interior of the switching curve, Sensor 1 is queried due to its low cost. Outside the switching curve, the estimation error is large enough to necessitate querying Sensor 2, which has higher precision.

Chapter 5

Optimal Transmission in a Time-Varying Channel: Heavy Traffic Analysis

5.1 Introduction

With the widespread deployment of wireless and ad-hoc networks, the energy-efficiency of wireless transmission in a fading channel has attracted much attention. It is now well understood that a transmission scheme that takes advantage of the time-varying character of a channel can significantly improve the use of scarce energy resources. As an extreme case, the policy that transmits only when the channel is in the best state can achieve the best energy efficiency while resulting in arbitrary long delay. Thus, there is clearly a trade-off between energy efficiency and delay constraints.

The problem of energy-efficient scheduling over a fading wireless channel has been studied under different delay constraints in the recent past [11, 38, 100]. In [38,100], the authors consider scheduling under a hard delay constraint, and maximize the throughput given energy and timing constraints. In [38], a finite horizon stochastic control formulation is used and a closed form solution to the dynamic programming equation is derived in some simplified cases. Berry and Gallager consider power control with delay constraints in an asymptotic sense [11]. They consider a single queue served by a fading channel. For a given data-arrival rate, the minimum power required to stabilize the queue can be computed directly from the capacity of the channel. However, with this minimum power, it is well known from queueing theory that the associated queueing delay is unbounded. The authors in [11] allocate an excess power ΔP and study the associated mean queuing delay D. They show that the optimal power control policy which takes both the channel state and the queue length into account results in an excess-power versus delay trade-off that behaves asymptotically as $\Delta P \propto \frac{1}{D^2}$. Further, they show that a single queue-length based threshold type policy achieves the same decay rate as the optimal policy, or in other words, the threshold policy is *order optimal* (however, they do not show optimality of the threshold policy).

In this chapter, we show optimality of the threshold policy in the so-call *heavy traffic regime*, or large queue asymptotics, for a single queue with a time-varying channel having a finite number of channel states indexed by $j \in \{1, 2, \ldots, N\}$. "Heavy traffic" is certain critical phase in a queueing system that arrival rates are close to service rates and queueing delay becomes large. Specifically, by proper scaling of the difference between arrival and service rates, the queueing dynamics can be approximated by a reflected diffusion process under mild conditions on the arrival and service processes. We impose both a peak power constraint $p_{\max}$, as well as an average power constraint $\bar{p}$ for power allocation, and the problem can be transformed into an ergodic control problem of reflected diffusion processes.

We work with the heavy-traffic limit for such a system under a fast channel variation assumption [2, 19, 21], whose dynamics are governed by a reflected Itô stochastic differential equation. We consider the problem of minimizing the long-term average value of a function $c(x)$ which depends on the heavy-traffic queue-length process x. The cost function $c(x)$ satisfies either *(i)* $c(x)$ is strictly increasing and bounded, or *(ii)* $c(x)$ grows unbounded (i.e., $c(x) \to \infty$, as $x \to \infty$). For example, $c(x) = x$ corresponds to minimizing the average queue length (or equivalently, from Little's law, the mean delay).

The main contributions of our work are:

(i) We show that when c is monotone, then the optimal control that minimizes the long-term average cost subject to the power constraints is a *channel state based threshold policy*. Specifically, associated with each channel state j there is a queue-threshold $\hat{x}_j$, such that at any time t, the optimal policy transmits at *peak power* $p_{\max}$ over channel state j, if the queue length $x(t) > \hat{x}_j$, and does not transmit otherwise. Further, using Lagrange duality and exploiting the monotonicity property of c, we reduce the problem of determining the queue-thresholds $\{\hat{x}_j, j = 1, 2, \ldots, N\}$ to that of solving a set of algebraic equations. Throughout the analysis we strive not to rely as much on the one-dimensional (one queue) character of the problem, aiming to present an approach that can scale up to higher dimensions.

(ii) An interpretation of the heavy-traffic limit is the following: Given a data arrival rate, sufficient "equilibrium" power is first allocated such that the capacity of the channel matches the arrival rate. Then, an amount of *excess power* is allocated based on the channel state and queue length. With such an interpretation, a special case of our result when the equilibrium power is allocated according to channel state dependent water-filling [26] (and is strictly positive in each channel state), results in the queue-length threshold being channel state invariant. In other words, for any monotone cost function $c(x)$, we have $\hat{x}_j = \hat{x}$, independent of channel state j. Thus, by applying the cost function $c(x) = x$, in this special case, our results indicate that the single-threshold policy derived in [11] is in fact asymptotically optimal.

(iii) For a system not in heavy-traffic, we numerically compute the optimal policy using dynamic programming, and compare this with the threshold policy that is optimal in the heavy-traffic limit. These numerical results indicate that the threshold policy performs close to the optimal policy even when the system is not in heavy-traffic.

(iv) From a technical standpoint, this problem falls under the domain of ergodic control of diffusions with constraints, and we adopt the convex analytic approach of [16,18]. The approach in [16] requires *both* the cost function as well as the constraint function (due to power constraints) to satisfy a *near-monotone* condition (see (5.17)). However, the constraint function is *not* near-monotone in our problem. Hence, since the results in [16] cannot be quoted, we first establish the existence of an optimal control within the class of stationary feedback controls. Next, using classical Lagrange multiplier theory, we show that the constrained problem is equivalent to an unconstrained one, namely minimizing the ergodic cost of the associated Lagrangian. We accomplish this by establishing that the near-monotone condition is satisfied for the Lagrangian (this result uses only the near-monotonicity of the cost function), and proceed to characterize the optimal policy for the unconstrained problem via the associated Hamilton Jacobi Bellman (HJB) equation. The solution to the original problem is then obtained by a straightforward application of Lagrange duality. We exhibit the structure of the optimal policy, and also establish that optimality holds over all non-anticipative policies, and not only over the stationary ones.

This chapter is organized as follows. In Section 5.2, the background and related work of heavy traffic analysis are introduced. Section 5.3 presents the Markovian model and the heavy-traffic model for the time-varying channel. In Section 5.4 we describe the optimal control problem and prove the existence of an optimal policy among stationary ones. In Section 5.5 we introduce the equivalent unconstrained problem using Lagrange multiplier theory and characterize the ergodic control problem relative to the Lagrangian via the HJB equation. We also show that the optimal policy has a multi-threshold structure. In Section 5.6 we present an analytical solution of the HJB equation. In order to demonstrate the approach, we specialize to the problem of minimizing the mean delay, i.e., $c(x) = x$, and derive closed form expressions for one and two-state channels. In Section 5.7, we evaluate the performance of the optimal policy for the heavy-traffic model by applying it to a system which does not operate in the heavy-traffic region.

5.2 Heavy traffic analysis: background and related work

In recent years, the heavy-traffic approximation has been successfully applied to performance evaluation and control of communication networks [5, 20, 21, 44, 46, 59, 79, 91, 94]. By heavy-traffic, we mean that the average fraction of time at which the server is free is small, or equivalently, the communication channel has little spare capacity. Largely due to this "small idle time" assumption, suitably scaled queueing processes can be well approximated by a reflected diffusion process. The communication networks are large in the sense that there are many users and the channel capacity is large. This size parameterizes the system and the scaling is called "fast arrivals/service" scaling. Other types of queueing systems are large in the sense that the small idle time implies that the queue size is large. Then with suitable scaling, the reflected diffusion approximation can be obtained, again under broad conditions. This scaling is called "large queue asymptotics".

The classical central limit theorem (CLT) can provide useful and accurate estimates even if the number of random variables involved is not very large, depending on their distributions. Based on a functional version of CLT, the reflected diffusion approximation to a network, which is essentially a second order approximation, can also provide accurate estimates even for "moderate" traffic. Simulations over a wide variety of practical problems and operating conditions show that the diffusion model yields very accurate approximation.

The benefits of the diffusion approximation mainly come from two aspects:

- Firstly, in the heavy traffic model, the state space is "aggregated", which can simplify the analysis and possibly yield structural results on the solution; and moreover, the "heavily congested" system can result in significant reduction on the model, e.g., state space collapse [20, 44, 94].

- Secondly, numerical methods (e.g. QNET [28, 29, 47]) are available to get the first and second moments of the stationary distributions for large scale systems if the model is uncontrolled and non-state-dependent, or even compute optimal controls if the dimension is not large [61].

Another view of the heavy traffic model is from the time-scale of the system. The randomness of arrivals and departures in the queueing system can be smoothed by looking at a larger time duration or faster time-scale. Intuitively, considering arrivals/departures within a large time-scale in the order of $O(n)$ with n large and scale the queue size by a factor $\frac{1}{n}$, the system can be approximated by the first-order model, or *fluid model* and we say its time-scale is time-scale $O(n)$. On the other hand, with arrival rate approaching the service rate in the order of $O(\sqrt{n})$, the heavy traffic model scales the queue size by a factor $\frac{1}{\sqrt{n}}$ and has a time-scale $O(\sqrt{n})$. Thus in the classic heavy traffic model, there are two time-scales: the real time-scale, where the empirical model lives; and the diffusion time-scale, where the limiting model stays. With more and more practical applications having time-varying factors and a changing environment, it is very important to incorporate the time-scale of time-varying parameters into the heavy traffic model.

To our best knowledge, the work by Buche and Kushner [21] appears to be the first work on heavy traffic analysis applied to systems with time-varying parameters. They apply the heavy-traffic approximation to model the multi-user power allocation problem in wireless fading channels, and design an optimal control in the heavy-traffic region. They consider the scenario where a fixed amount of power is available at each time slot, and this power needs to be allocated to multiple users according to their queue length and current channel states. There are three time-scales involved in the model: the real time-scale, the channel-fading time scale, and the diffusion time scale. They show that the optimal policy is a switching curve by numerical results.

In the rest of this chapter, we adopt a similar scaling for the time varying channels as Buche and Kushner in [21] and derive the optimal power allocation scheme for a wireless fading channel under the heavy traffic approximation. The solution is given in a close-form and a threshold policy is shown to be optimal in the heavy traffic regime.

5.3 The system model and the heavy-traffic limit

We consider a queuing system that consists of a transmitter operating over a fading channel (see Figure 5.1). Time is assumed to be divided into discrete slots, and the channel state process is an irreducible, aperiodic, finite state Markov chain $L(t)$ with N states having a stationary distribution $\pi = (\pi_1, \dots, \pi_N)$. The channel gain is denoted by g_j when the channel state $L(t) = j$, and the power P allocated at time t determines the service rate $r(P, j)$ of the queue. For example, given the power P, bandwidth W and channel gain g_j, $r(P, j) = W \log_2(1 + Pg_j)$ is the Shannon capacity, the upper bound of the channel transmission rate. The service rate $r(P, j)$ can take different forms for practical systems depending on the details of modulation and coding.

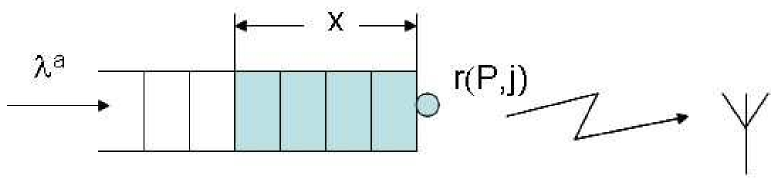


Figure 5.1: A transmitter sends packets to a receiver through a time-varying wireless channel.

As is common in heavy-traffic analysis, we construct a sequence of queueing systems indexed by n, such that as $n \to \infty$, the transmitter idle time goes to zero in an appropriate manner (see (5.1) below). In the heavy-traffic approximation, there are two time scales: one is the time scale the real system works on; the other is the diffusion time scale, which is a slower scale. A small time period Δt in the diffusion time scale contains a large number of arrivals and departures, which is of order $\mathcal{O}(n\Delta t)$. For a wireless channel with time-varying characteristics, there is yet another time scale, i.e., the time scale of channel variation. We consider the fast channel variation model [2,19], which assumes that the channel variation has a time scale faster than the diffusion time scale, but slower than the arrival process time scale, as shown in Figure 5.2. Thus, for the n-scaled queueing system, the channel process is $L(n^{-\kappa}t)$, where $\kappa \in (0,1)$. As a result, over an interval of time $n\Delta t$, the number of channel transitions is $\mathcal{O}(n^{1-\kappa}\Delta t)$, and the number of arrivals *within each channel state* (i.e, between any pair of channel transitions) is $\mathcal{O}(n^{\kappa}\Delta t)$. Thus, the total number of arrivals over the time interval $n\Delta t$ is $\mathcal{O}(n\Delta t)$.

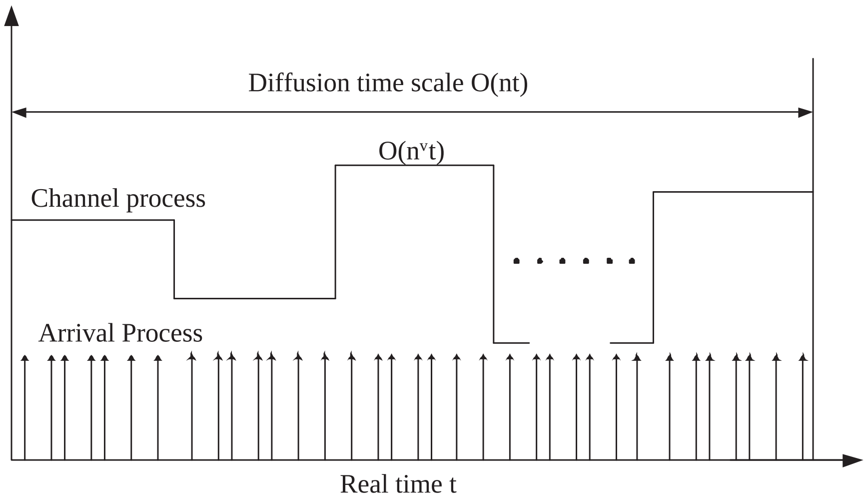


Figure 5.2: The three time scales of the heavy-traffic model under the fast channel variation assumption.

Practically, this scaling fits into the scenario that the channel changes slowly compared to the packet arrival rate, i.e., a slowly fading channel such as an indoor wireless environment, or a low-mobile-velocity outdoor wireless environment [82]. For instance, with 1xEV-DO (the 3G wireless data service), a scheduling time-slot is 1.667 msec, which corresponds to the arrival time-scale. For a mobile user with velocity 6 mph, the channel coherence time, which corresponds to the time-scale of channel changes, is about 50 msec. Thus, the scaling we use here seems applicable in these practical regimes.

We consider a sequence of queueing systems indexed by n, with the queue length $x^n(t)$, arrival process $A^n(t)$ and departure process $D^n(t)$, which can be controlled by transmission power. For the queueing system indexed by n, we denote the l-th inter-arrival time by ζ_l^n, and assume it satisfies the following assumption [59].

Assumption 5.3.1. *The inter-arrival intervals $\{\zeta_l^n,\ l \in \mathbb{N}\}$ satisfy the following:*

1. $\left\{|\zeta_l^n|^2,\ l, n \in \mathbb{N}\right\}$ *is uniformly integrable.*
2. *For each n, $\{\zeta_l^n,\ l \in \mathbb{N}\}$ are independent. Moreover, there exist constants $\bar{\zeta}^n$, $\bar{\zeta}$, σ_a^2, such that*

$$\mathbb{E}[\zeta_l^n] = \bar{\zeta}^n \xrightarrow[n\to\infty]{} \bar{\zeta}\,, \qquad \lim_{n\to\infty} \mathbb{E}\left[1 - \frac{\zeta_l^n}{\bar{\zeta}^n}\right]^2 = \sigma_a^2\,.$$

3. *The inter-arrivals are independent of the channel process.*

Note that if either ζ_l^n are identically distributed with finite variance, or ζ_l^n are deterministic but periodic, Assumption 5.3.1 is satisfied. The mean arrival rate for the n-th system is defined as $\lambda_n^a = 1/\bar{\zeta}^n$ and the limiting arrival rate λ^a is defined as $\lambda^a = 1/\bar{\zeta}$.

For the queue indexed by n, the service rate r is controlled by the transmission power P_n. Under the heavy-traffic approximation, we suppose that mean arrival rate converges to the service rate under the scaling,

$$\lim_{n\to\infty} \Big(\lambda_n^a - \mathbb{E}\big[r(P_n)\big]\Big) n^{\frac{1-\kappa}{2}} = \text{constant}\,. \tag{5.1}$$

for some $\kappa \in (0, 1)$. Assuming (5.1) holds, we decompose the power allocation $P(q, j)$ for buffer size q, and channel state j into

$$P_n(q, j) = P_0(j) + n^{-\frac{1-\kappa}{2}} u_j(q)\,.$$

The "equilibrium" power $P_0(j)$ is allocated in such a manner that

$$\lambda^a = \sum_{j=1} r(P_0(j), j)\pi_j \,, \tag{5.2}$$

Remark 5.3.1. *Note that the optimal allocation of the equilibrium power gives rise to a static optimization problem, namely, minimize the average power $\mathbb{E}[P]$ given the service rate $\mathbb{E}[r(P)] \geq \lambda^a$, where $\mathbb{E}[\,\cdot\,]$ is taken over the channel distribution. For a fading channel with additive white Gaussian noise (AWGN), water-filling is the optimal way for allocating power subject to* (5.2) *in an information theoretic sense [26]. In general, the equilibrium allocation can be computed numerically.*

We assume that the equilibrium power has been allocated, either by water-filling or by numerically determining the optimal allocation, and we address the problem of optimally allocating the residual power. Optimality here is in an asymptotic sense, i.e., pertains to the limiting system under heavy-traffic conditions. By expanding the service rate $r(P, j)$ around $P = P_0(j)$, using Taylor's series, we obtain

$$r(P, j) = r(P_0(j), j) + \frac{u_j}{n^{\frac{1-\kappa}{2}}} \frac{\partial r}{\partial P}(P_0(j), j) + o(n^{-\frac{1-\kappa}{2}})\,.$$

Let

$$r_0(j) := r(P_0(j), j)\,, \qquad \gamma_j := \frac{\partial r}{\partial P}(P_0(j), j)\,.$$

Then

$$r(P, j) \approx r_0(j) + n^{-\frac{1-\kappa}{2}} \gamma_j u_j\,. \tag{5.3}$$

Thus, $\lambda^a = \sum_{j=1} r_0(j)\pi_j$, and the incremental service rate gained from the residual amount of power u is $n^{-\frac{1-\kappa}{2}} b(u)$, where

$$b(u) = \sum_{j=1}^{N} \gamma_j \pi_j u_j\,.$$

Remark 5.3.2. *We observe that if the equilibrium power $\{P_0(j)\}$ is allocated according to channel-state dependent water-filling [26], and if such an allocation results in $P_0(j) > 0$ for all channel states j, then $\gamma_i = \gamma_j$ for all i, j.*

Next, defining $x^n(t) := n^{-\frac{(1+\kappa)}{2}} q(nt)$ and using the techniques in [21], we show in Appendix A.1 that $x^n(t)$ converges weakly to a limiting queueing system as $n \to \infty$. The dynamics of the limiting queueing system are governed by the equation

$$x(t) = x(0) - \int_0^t b\big(u(s)\big)\, ds + \sigma W(t) + z(t)\,, \tag{5.4}$$

where $x(t)$ is the queue-length process, $W(t)$ is the standard Wiener process, σ is a positive constant, $z(t)$ is a nonincreasing process and grows only at those points t for which $x(t) = 0$, and $x(t) \geq 0$, for all $t \geq 0$. The process $z(t)$, which ensures that the queue-length $x(t)$ remains non-negative, is uniquely defined. For further details see [23, pp. 128, Theorem 6.1] and [41, pg. 178]. The corresponding Itô stochastic differential equation describing the heavy-traffic dynamics takes the form

$$dx(t) = -b\big(u(t)\big)dt + \sigma dW(t) + dz(t)\,. \tag{5.5}$$

5.4 The optimal control problem for the heavy-traffic Model

The optimization problem of interest for the non-scaled queueing system is to minimize (pathwise, a.s.) the long-term average queueing length (and thus, from Little's law, the mean delay)

$$\limsup_{T\to\infty} \frac{1}{T} \int_0^T q(t)\, dt\,,$$

or more generally, to minimize the long-term average value of some penalty function $c : \mathbb{R}_+ \to \mathbb{R}$, i.e.,

$$\limsup_{T\to\infty} \frac{1}{T} \int_0^T c\big(q(t)\big)\, dt\,,$$

subject to a constraint on the average available power of the form

$$\limsup_{T\to\infty} \frac{1}{T} \int_0^T P\big(q(t), L(t)\big)\, dt \leq P_{\text{avg}}\,.$$

It is well known from queueing theory, that if only the basic power P_0 is allocated, which matches the service rate to the arrival rate, then the resulting traffic intensity is equal to 1, and the queueing delay diverges. However, choosing the control term u appropriately can result in a bounded average queue length. In the heavy-traffic model described in Section 5.3, once the channel model is provided, v is fixed, and only the excess power u can be used to control the queue. Thus the original optimization problem transforms to an analogous problem in the limiting system, namely,

$$\text{minimize} \qquad \limsup_{T\to\infty} \frac{1}{T}\int_0^T c\bigl(x(t)\bigr)\,dt\,, \quad \text{a.s.} \tag{5.6a}$$

$$\text{subject to} \qquad \limsup_{T\to\infty} \frac{1}{T}\int_0^T h\bigl(u(t)\bigr)\,dt \le \bar{p}\,, \quad \text{a.s.} \tag{5.6b}$$

where

$$h(u) = h(u_1,\dots,u_N) = \sum_{j=1}^{N} \pi_j u_j\,.$$

The control variable u takes values in $U := \bigl[0, p_{\max}\bigr]^N$, with $p_{\max}$ denoting the (excess) peak power, and $\bar{p}$ denoting the (excess) average power. Naturally, for the constraint in (5.6b) to be feasible $\bar{p} \le p_{\max}$.

Definition 5.4.1. *Let $\mathfrak{F}_t := \sigma\{W(s),\ s \le t\}$. The minimization in* (5.6) *is over all control processes $u(t)$ which are $\{\mathfrak{F}_t\}$-adapted, have measurable sample paths and are* non-anticipative*: for $s \le t$, $W(t) - W(s)$, and $\sigma\bigl\{u(r), W(r),\ r \le s\bigr\}$ are independent. Such a process u is called an* admissible control *and the class of admissible controls is denoted by $\mathfrak{U}$. An admissible control which takes the form $u(t) = v\bigl(x(t)\bigr)$, for some measurable function $v : \mathbb{R}_+ \to U$ is called a* stationary (Markov) control*, and we denote this class by $\mathfrak{U}_s$.*

Given a measurable function $v : \mathbb{R}_+ \to U$, the stochastic differential equation in (5.5) under the control $u(t) = v\bigl(x(t)\bigr)$ has a unique strong solution, which is a Feller-Markov process. Let $\mathbb{E}^v_x$ denote the expectation operator on the path space of the process, with initial condition $x(0) = x$, and T^v_t denote the Markov semigroup acting on the space of bounded continuous functions $\mathcal{C}_b(\mathbb{R}+)$, defined by $T^v_t f(x) = \mathbb{E}^v_x\bigl[f\bigl(x(t)\bigr)\bigr]$, $f \in \mathcal{C}_b(\mathbb{R}+)$. It is known that T^v_t has infinitesimal generator $\mathcal{L}^{v(x)}$

(see [36, pg. 366-367], [37]), where

$$\mathcal{L}^u := \frac{\sigma^2}{2}\frac{d^2}{dx^2} - b(u)\frac{d}{dx}\,, \quad u \in U\,.$$

The boundary at 0, imposes restrictions on the domain of $\mathcal{L}^u$ (see [36, pg. 366-367]).

The generator $\mathcal{L}$ can be readily used to compute functionals of the process. As asserted in [37, pg. 80], if f is a bounded measurable function on $\mathbb{R}_+$ then $\varphi(x,t) = \mathbb{E}_x^v\bigl[f\bigl(x(t)\bigr)\bigr]$ is a generalized solution of the problem

$$\begin{aligned} \frac{\partial\varphi}{\partial t}(x,t) = \mathcal{L}^v\varphi(x,t)\,, \quad x \in (0,\infty)\,,\ t > 0\,, \\ \varphi(x,0) = f(x)\,, \qquad \frac{\partial\varphi}{\partial x}(0,t) = 0\,. \end{aligned} \tag{5.7}$$

Also, Itô's formula can be applied as follows [54, pg. 500, Lemma 4], [58]: If $\varphi \in \mathcal{W}^{2,p}(\mathbb{R}_+)$ is a bounded function (here $\mathcal{W}$ stands for the Sobolev space) satisfying $\frac{d\varphi}{dx}(0) = 0$, then for $t \geq 0$,

$$\mathbb{E}_x^v\bigl[\varphi\bigl(x(t)\bigr)\bigr] - \varphi(x) = \mathbb{E}_x^v\biggl[\int_0^t \mathcal{L}^v\varphi\bigl(x(t)\bigr)\,dt\biggr]\,. \tag{5.8}$$

Definition 5.4.2. *A control $v \in \mathfrak{U}_s$ is called* stable *if the resulting $x(t)$ is positive recurrent. We denote the class of stable controls by $\mathfrak{U}_{ss}$. A control $v \in \mathfrak{U}_s$ is called* bang-bang*, or* extreme*, if $v(x) \in \{0, p_{max}\}^N$, for almost all $x \in \mathbb{R}_+$. We refer to the class of extreme controls in $\mathfrak{U}_{ss}$ as* stable extreme *controls and denote it by $\mathfrak{U}_{se}$.*

Let $\mathcal{P}(\mathbb{R}_+)$ denote the set of probability measures on the Borel σ-field of $\mathbb{R}_+$. Recall that a probability measure $\mu \in \mathcal{P}(R_+)$ is said to be invariant for process $x(t)$ under the control $v \in \mathfrak{U}_\mathrm{s}$, if $\int T_t^v f\,d\mu = \int f\,d\mu$, for all $f \in \mathcal{C}_b(R_+)$, and $t \geq 0$. It is the case that if $v \in \mathfrak{U}_\mathrm{ss}$, then the controlled process $x(t)$ has a unique invariant probability measure μ_v which is absolutely continuous with respect to the Lebesgue measure. Let $\mathcal{C}_c^\infty(0,\infty)$ denote the class of smooth functions in $(0,\infty)$ with compact support. We make frequent use of the following characterization. A necessary and sufficient condition for a probability measure $\mu \in \mathcal{P}(R_+)$ to be an invariant probability measure of the controlled process $x(t)$ under $v \in \mathfrak{U}_{\mathrm{s}}$ is

$$\int_{\mathbb{R}_+} \mathcal{L}^v g(x)\, \mu(dx) = 0\,, \quad \forall g \in \mathcal{C}_c^\infty(0,\infty)\,. \tag{5.9}$$

Necessity of (5.9) is a straightforward application of (5.8) and the definition of an invariant measure. Borkar establishes sufficiency for diffusions without reflection, by employing the uniqueness of the Cauchy problem for the forward Kolmogorov equation [15, pg. 144, Lemma 1.2]. The boundary complicates matters for this approach, so we employ the following result, which we state in the d-dimensional setting. Let $D \subset \mathbb{R}^d$ be a domain and $\mathcal{L}$ a second order uniformly elliptic operator with bounded measurable coefficients in D, and with the second order coefficients Lipschitz continuous. If μ is a finite Borel measure on D satisfying $\int_D \mathcal{L}g(x)\,\mu(dx) = 0$, for all $g \in \mathcal{C}_c^\infty(D)$, then μ is absolutely continuous with respect to the Lebesgue measure, i.e., has density [13, Theorem 2.1]. Thus, if μ satisfies (5.9), then $\mu(dx) = f_v(x)\,dx$, and hence using the adjoint operator $(\mathcal{L}^v)^*$ we have

$$\int_{\mathbb{R}_+} g(x)(\mathcal{L}^v)^* f_v(x)\,dx = 0\,, \quad \forall g \in \mathcal{C}_c^\infty(0,\infty)\,,$$

which is equivalent to $(\mathcal{L}^v)^* f_v = 0$. Following the proof of [9, pg. 87, Proposition 8.2] and utilizing (5.7), we deduce that f_v is indeed the density of an invariant probability distribution. It follows from the preceding discussion that f_v is the density of an invariant probability measure μ_v if and only if it is a solution of the Fokker-Planck equation

$$\bigl(\mathcal{L}^v\bigr)^* f_v(x) = \frac{d}{dx}\left(\frac{\sigma^2}{2}\frac{df_v}{dx}(x) + b\bigl(v(x)\bigr) f_v(x)\right) = 0\,. \tag{5.10}$$

Moreover, solving (5.10), we deduce that $v \in \mathfrak{U}_{\mathrm{s}}$ is stable if and only if

$$A_v := \int_0^\infty \exp\biggl(-\frac{2}{\sigma^2}\int_0^x b\bigl(v(y)\bigr)\,dy\biggr)\,dx < \infty\,,$$

in which case the solution of (5.10) takes the form

$$f_v(x) = A_v^{-1} \exp\biggl(-\frac{2}{\sigma^2}\int_0^x b\bigl(v(y)\bigr)\,dy\biggr)\,. \tag{5.11}$$

We work under the assumption that c has the following monotone property:

Table 5.1: Table of Symbols

Symbol	**Definition**	**First Appearance**
$\mathfrak{U}$ ($\mathfrak{U}_{\mathrm{s}}$)	admissible (stationary) controls	Def. 5.4.1
$\mathfrak{U}_{\mathrm{ss}}$ ($\mathfrak{U}_{\mathrm{se}}$)	stable stationary (extreme) controls	Def. 5.4.2
$\mathcal{P}(X)$	probability measures on X	Sec. 5.4
$\mathcal{G}$	set of ergodic occupation measures	Sec. 5.4.1
$\mathcal{M}$	set of invariant probability measures	Sec. 5.4.1
$H(\bar{p})$	subset of $\mathcal{G}$ with power bound $\bar{p}$	(5.13)

Assumption 5.4.1. *The function c is continuous and either it is asymptotically unbounded, i.e.,* $\liminf_{x\to\infty}\, c(x) = \infty$*, or if c is bounded then it is strictly increasing. In the latter case we define*

$$c_\infty := \lim_{x\to\infty}\, c(x)\,.$$

The analysis and solution of the optimization problem proceeds as follows: We first show that optimality is achieved for (5.6) relative to the class of stationary controls. Next, in Section 5.5 using the theory of Lagrange multipliers we formulate an equivalent unconstrained optimization problem. We show that an optimal control for the unconstrained problem can be characterized via the HJB equation. This accomplishes two tasks. First, it enables us to study the structure of the optimal policies. Second, we show that this control is optimal among all controls in $\mathfrak{U}$. An analytical solution of the HJB equation is presented in Section 5.6. A list of symbols is included in Table. 5.1 for quick reference.

5.4.1 Existence of optimal stationary controls

In this subsection, we show that if the optimization problem in (5.6) is restricted to stationary controls, then there exists $v \in \mathfrak{U}_{\mathrm{se}}$ which is optimal.

Due to the presence of the constraint in (5.6b), the study of the optimization problem in (5.6) is more amenable by convex analytic arguments. We follow the approach in [16, 18]. However, we take advantage of the fact that the set of power levels U is convex and avoid transforming the problem to the relaxed control framework. Instead, we view U as the space of product probability measures on

$\{0, p_{\max}\}^N$. This is simply stating that for each j, u_j may be represented as a convex combination of the '0' power-level and the peak power $p_{\max}$. In other words, U is viewed as a space of relaxed controls relative to the discrete control input space $\{0, p_{\max}\}^N$. This has the following advantage: by showing that optimality is attained in the set of precise controls, we assert the existence of a control in $\mathfrak{U}_{\mathrm{se}}$ which is optimal. Another important point is that the convex analytic method in [16, 18] for the constrained problem is not equipped to establish optimality of a stationary policy over all admissible controls. This issue is dealt with in Section 5.5, and is resolved by employing the HJB equation.

Let $\mathcal{M} \subset \mathcal{P}(\mathbb{R}_+)$ denote the set of all invariant probability measures μ_v of the process $x(t)$ under the controls $v \in \mathfrak{U}_{\mathrm{ss}}$. Let $\widetilde{U} := \{0, p_{\max}\}^N$. The generic element of $\widetilde{U}$ takes the form $\tilde{u} = (\tilde{u}_1, \dots, \tilde{u}_N)$, with $\tilde{u}_i \in \{0, p_{\max}\}$, $i = 1, \dots, N$. There is a natural isomorphism between U and the space of product probability measures on $\widetilde{U}$ which we denote by $\mathcal{P}_{\otimes}(\widetilde{U})$. This is viewed as follows. Let δ_p denote the Dirac probability measure concentrated at $p \in \mathbb{R}_+$. For $u \in U$, we associate the probability measure $\tilde{\eta}_u \in \mathcal{P}_{\otimes}(\widetilde{U})$ defined by

$$\tilde{\eta}_u(\tilde{u}) := \bigotimes_{i=1}^{N} \left[\left(1 - \frac{u_i}{p_{\max}}\right) \delta_0(\tilde{u}_i) + \frac{u_i}{p_{\max}} \delta_{p_{\max}}(\tilde{u}_i) \right],$$

for $\tilde{u} \in \widetilde{U}$. Similarly, given $v \in \mathfrak{U}_{\mathrm{ss}}$ we define $\eta_v : \mathbb{R}_+ \to \mathcal{P}_{\otimes}(\widetilde{U})$ and and $\nu_v \in \mathcal{P}(\mathbb{R}_+ \times \widetilde{U})$ by

$$\begin{aligned} \eta_v(x, d\tilde{u}) &:= \tilde{\eta}_{v(x)}(d\tilde{u}) \\ \nu_v(dx, d\tilde{u}) &:= \mu_v(dx)\eta_v(x, d\tilde{u}), \end{aligned}$$

where $\mu_v \in \mathcal{M}$ is the invariant probability measure of the process under the control $v \in \mathfrak{U}_{\mathrm{ss}}$. The set of *ergodic occupation* measures is defined as $\mathcal{G} := \{\nu_v : v \in \mathfrak{U}_{\mathrm{ss}}\}$. It follows by (5.9) that $\nu \in \mathcal{G}$ if and only if

$$\int_{\mathbb{R}_+} \mathcal{L}^{\tilde{u}} g(x)\, \nu(dx, d\tilde{u}) = 0, \quad \forall g \in \mathcal{C}_c^\infty(0, \infty). \tag{5.12}$$

Due to the linearity of $u \mapsto h(u)$, we have the following identity (which we choose to express as an integral rather than a sum, despite the fact that $\widetilde{U}$ is a finite space):

$$h\big(v(x)\big) = \int_{\widetilde{U}} h(\tilde{u})\eta_v(x, d\tilde{u})\,, \quad v \in \mathfrak{U}_{\mathrm{ss}}\,,$$

As a point of clarification, 'h' inside this integral is interpreted as the restriction of h on $\widetilde{U}$. The analogous identity holds for $b(u)$.

In this manner we have defined a model whose input space $\widetilde{U}$ is discrete, and for which the original input space U provides an appropriate convexification. Note however that $U \sim \mathcal{P}_{\otimes}(\widetilde{U})$ is not the input space corresponding to the relaxed controls based on $\widetilde{U}$. The latter is $\mathcal{P}(\widetilde{U})$, which is isomorphic to a 2^N-simplex in $\mathbb{R}^{2^N-1}$, whereas $\mathcal{P}_{\otimes}(\widetilde{U})$ is isomorphic to a cube in $\mathbb{R}^N$. We select $\mathcal{P}_{\otimes}(\widetilde{U})$ as the input space mainly because it is isomorphic to U. Since there is a one to one correspondence between the extreme points of $\mathcal{P}_{\otimes}(\widetilde{U})$ and $\mathcal{P}(\widetilde{U})$, had we chosen to use the latter, the analysis and results would have remained unchanged. Even though we are not using the standard relaxed control setting, since $\mathcal{P}_{\otimes}(\widetilde{U})$ is closed under convex combinations and limits, the theory goes through without any essential modifications.

For $\bar{p} \in (0, p_{\max}]$, let

$$H(\bar{p}) := \left\{\nu \in \mathcal{G} : \int_{\mathbb{R}_+ \times \widetilde{U}} h(\tilde{u})\,\nu(dx, d\tilde{u}) \le \bar{p}\right\}\,. \tag{5.13}$$

Then $H(\bar{p})$ is a closed, convex subset of $\mathcal{G}$. It is easy to see that it is also nonempty, provided $\bar{p} > 0$. Indeed, let $x' \in \mathbb{R}_+$ and consider the policy $v_{x'}$ defined by

$$(v_{x'})_i = \begin{cases} 0\,, & x \le x' \\ p_{\max}\,, & x > x'\,, \end{cases} \qquad i = 1, \dots, N\,.$$

Under this policy, the diffusion process in (5.5) is positive recurrent and its invariant probability measure has a density $f_{x'}$ which is a solution of (5.10). Let

$$\alpha_k := \frac{2p_{\max}}{\sigma^2} \sum_{i=1}^{k} \gamma_i \pi_i\,, \quad k = 1, \dots, N\,. \tag{5.14}$$

The solution of (5.10) takes the form

$$f_{x'}(x) = \frac{\alpha_N e^{-\alpha_N (x-x')^+}}{1+\alpha_N x'}\,.$$

where $(y)^+ := \max(y,0)$. Then

$$\int_{\mathbb{R}_+} h\bigl(v(x)\bigr) f_{x'}(x)\,dx = \frac{p_{\max}}{1+\alpha_N x'}\,,$$

and it follows that $\nu_{v_{x'}} \in H(\bar{p})$, provided

$$x' \geq \frac{1}{\alpha_N}\Bigl(\frac{p_{\max}}{\bar{p}} - 1\Bigr)\,.$$

Thus, the optimization problem in (5.6) when restricted to stationary, stable controls is equivalent to

$$\text{minimize over } \nu \in H(\bar{p}) : \quad \int_{\mathbb{R}_+ \times \widetilde{U}} c(x)\,\nu(dx, d\tilde{u})\,. \tag{5.15}$$

We also define

$$J^*(\bar{p}) := \inf_{\nu \in H(\bar{p})} \int_{\mathbb{R}_+ \times \widetilde{U}} c\,d\nu\,. \tag{5.16}$$

We proceed as follows. It is well known that $\mathcal{G}$ and $\mathcal{M}$ are convex and that their extreme points $\mathcal{G}_e$ and $\mathcal{M}_e$ correspond to controls in $\mathfrak{U}_{\mathrm{se}}$. It is shown in [16,18] that, under a near-monotone assumption on both the running cost c and h the infimum in (5.16) is attained in $H(\bar{p})$. This near-monotone condition amounts to

$$\liminf_{x\to\infty}\ c(x) > J^*(\bar{p}) \tag{5.17a}$$

$$\inf_{\tilde{u}\in\widetilde{U}}\ h(\tilde{u}) > \bar{p}\,. \tag{5.17b}$$

Clearly (5.17b) does not hold, and hence the results in [16, 18] cannot be quoted to assert existence. So we show directly in Theorem 5.4.3 that (5.15) attains a minimum in $H(\bar{p})$, and more specifically that this minimum is attained in $\mathfrak{U}_{\mathrm{se}}$.

Concerning the extreme points of $\mathcal{G}$, the following lemma is a variation of [18, Lemma 3.5].

Lemma 5.4.1. *Let $A \subset \mathbb{R}_+$ be a bounded Borel set of positive Lebesgue measure. Suppose that v', $v'' \in \mathfrak{U}_s$ differ a.e. on A and agree on A^c, and that for some $v_0 \in \mathfrak{U}_{ss}$ and measurable $r : \mathbb{R}_+ \to [0,1]$, which satisfies $r(x) \in (0,1)$, for almost all $x \in A$, we have*

$$v_0(x) = r(x)v'(x) + \big(1 - r(x)\big)v''(x)\,. \tag{5.18}$$

Then, there exist $\hat{v}'$, $\hat{v}'' \in \mathfrak{U}_{ss}$ which differ a.e. on A and agree on A^c, such that

$$\nu_{v_0} = \tfrac{1}{2}(\nu_{\hat{v}'} + \nu_{\hat{v}''})\,.$$

In particular ν_{v_0} is not an extreme point of $\mathcal{G}$.

Since, every $v \in \mathfrak{U}_{\mathrm{ss}} \setminus \mathfrak{U}_{\mathrm{se}}$ can be decomposed as in (5.18) satisfying the hypotheses of Lemma 5.4.1, we obtain the following corollary.

Corollary 5.4.2. *If $\nu_v \in \mathcal{G}_e$ then $v \in \mathfrak{U}_{se}$.*

The main result of this section is contained in the following theorem whose proof can be found in Appendix A.2.

Theorem 5.4.3. *Under Assumption 5.4.1, for any $\bar{p} \in (0, p_{max}]$, there exists $v^* \in \mathfrak{U}_{se}$ such that ν_{v^*} attains the minimum in* (5.15)*.*

5.5 Lagrange multipliers and the HJB equation

In order to study the stationary optimal policies for (5.15), we introduce a parameterized family of unconstrained optimization problems that is equivalent to the problem in (5.6) in the sense that stationary optimal policies for the former are also optimal for the latter and vice-versa. We show that optimal policies for the unconstrained problem can be derived from the associated HJB equation. Hence, by studying the HJB equation we characterize the stationary optimal policies (5.15). We show that these are of a multi-threshold type and this enables us to reduce the optimal control problem to that of solving a system of $N + 1$ algebraic equations. Furthermore, we show that optimality is achieved over the class of all admissible policies $\mathfrak{U}$, and not only over $\mathfrak{U}_{\mathrm{s}}$.

With $\lambda \in \mathbb{R}_+$ playing the role of a Lagrange multiplier, we define

$$\begin{aligned} L(x,u,\bar{p},\lambda) &:= c(x) + \lambda\big(h(u) - \bar{p}\big) \\ \tilde{J}(v,\bar{p},\lambda) &:= \limsup_{T\to\infty} \frac{1}{T}\int_0^T L\big(x(t),v(t),\bar{p},\lambda\big)\,dt \\ \tilde{J}^*(\bar{p},\lambda) &:= \inf_{v\in\mathfrak{U}_{\mathrm{ss}}} \tilde{J}(v,\bar{p},\lambda)\,. \end{aligned} \tag{5.19}$$

The choice of the optimization problem in (5.19) is motivated by the fact that $J^*(\bar{p})$, defined in (5.16) is a convex, decreasing function of $\bar{p}$. This is rather simple to establish. Let $\bar{p}'$, $\bar{p}'' \in (0, p_{\max}]$ and denote by ν', ν'' the corresponding ergodic occupation measures that achieve the minimum in (5.15). Then, if $\delta \in [0,1]$, $\nu_0 := \delta\nu' + (1-\delta)\nu''$ satisfies $\int h\,d\nu_0 = \delta\bar{p}' + (1-\delta)\bar{p}''$, and since ν_0 is suboptimal for the optimization problem in (5.15) with power constraint $\delta\bar{p}' + (1-\delta)\bar{p}''$, we have

$$J^*(\delta\bar{p}' + (1-\delta)\bar{p}'') \le \int c\,d\nu_0 = \delta J^*(\bar{p}') + (1-\delta)J^*(\bar{p}'')\,.$$

A separating hyperplane which is tangent to the the graph of the function $J^*(\cdot)$ at a point $\big(\bar{p}_0, J^*(\bar{p}_0)\big)$, with $\bar{p}_0 \in (0, p_{\max}]$ takes the form

$$\big\{(\bar{p}, J) : J + \lambda_{\bar{p}_0}(\bar{p} - \bar{p}_0) = J^*(\bar{p}_0)\big\}\,,$$

for some $\lambda_{\bar{p}_0} \in \mathbb{R}_+$ (see Figure 5.3).

Standard Lagrange multiplier theory yields the following (see [67, pg. 217, Thm. 1]):

Theorem 5.5.1. *Let $\bar{p}_0 \in (0, p_{max}]$. There exists $\lambda_{\bar{p}_0} \in \mathbb{R}_+$, such that the minimization problem in* (5.15)*, over $H(\bar{p}_0)$ as well as the problem*

$$\textit{minimize}: \quad \int_{\mathbb{R}_+\times\tilde{U}} L(x,\tilde{u},\bar{p}_0,\lambda_{\bar{p}_0})\,\nu(dx,d\tilde{u}) \tag{5.20}$$

over $\nu \in \mathcal{G}$, both attain the same minimum value $J^*(\bar{p}_0) = \tilde{J}^*(\bar{p}_0, \lambda_{\bar{p}_0})$, at some $\nu_0 \in H(\bar{p}_0)$. In particular,

$$\int_{\mathbb{R}_+\times\tilde{U}} h(\tilde{u})\,\nu_0(dx,d\tilde{u}) = \bar{p}_0\,.$$

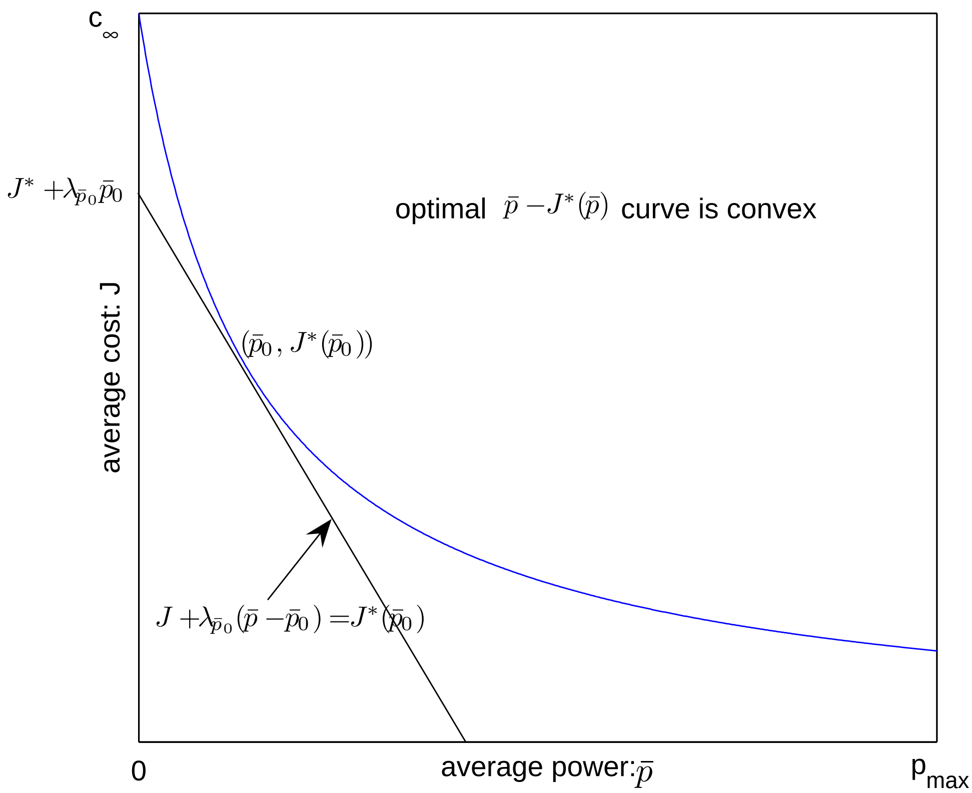


Figure 5.3: Convexity of $\bar{p} \mapsto J^*(\bar{p})$ and the separating hyperplane through $(\bar{p}_0, J^*(\bar{p}_0))$.

Characterizing the optimal policy via the HJB equation associated with the unconstrained problem in (5.20), is made possible by first showing that under Assumption 5.4.1 the cost $L(x, u, \bar{p}, \lambda)$ is near-monotone (see (5.22) below), and then employing the results in [15]. It is not difficult to show that under Assumption 5.4.1

$$\lim_{\bar{p}\to 0} J^*(\bar{p}) = \lim_{x\to\infty} c(x)\,. \tag{5.21}$$

Indeed, for $\bar{p} \in (0, p_{\max}]$, suppose $v \in \mathfrak{U}_{\mathrm{ss}}$ such that $\nu_v \in H(\bar{p})$. Letting $\gamma_{\max} := \max_i \{\gamma_i\}$, and using (5.11) we obtain

$$\begin{aligned} \bar{p} &\ge \int_0^\infty h\big(v(x)\big) f_v(x)\,dx \\ &\ge \gamma_{\max}^{-1} \int_0^\infty b\big(v(x)\big) f_v(x)\,dx \\ &= \frac{\sigma^2}{2\gamma_{\max} A_v}\,. \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^\infty c(x) f_v(x)\,dx &\ge \min_{x \ge \frac{1}{\sqrt{\bar{p}}}} \big\{c(x)\big\} \int_{\frac{1}{\sqrt{\bar{p}}}}^\infty f_v(x)\,dx \\ &= \min_{x \ge \frac{1}{\sqrt{\bar{p}}}} \big\{c(x)\big\} \left(1 - \frac{A_v^{-1}}{\sqrt{\bar{p}}}\right) \\ &\ge \min_{x \ge \frac{1}{\sqrt{\bar{p}}}} \big\{c(x)\big\} \left(1 - \frac{2\gamma_{\max}}{\sigma^2}\sqrt{\bar{p}}\right)\,. \end{aligned}$$

Hence,

$$J^*(\bar{p}) \ge \min_{x \ge \frac{1}{\sqrt{\bar{p}}}} \big\{c(x)\big\} \left(1 - \frac{2\gamma_{\max}}{\sigma^2}\sqrt{\bar{p}}\right)$$

and (5.21) follows. We need the following lemma, whose proof is contained in Appendix A.2.

Lemma 5.5.2. *Let Assumption 5.4.1 hold and suppose c is bounded. Then for any $\bar{p} \in (0, p_{max}]$, we have*

$$J^*\big(\tfrac{\bar{p}}{2}\big) < \tfrac{1}{2}\big(J^*(\bar{p}) + c_\infty\big)\,.$$

We are now ready to establish the near-monotone property of L. First, we introduce some new notation. For $\bar{p} \in (0, p_{\max}]$, let

$$\Lambda(\bar{p}) := \left\{\lambda \in \mathbb{R}_+ : J^*(\bar{p}') \geq J^*(\bar{p}) + \lambda(\bar{p} - \bar{p}')\,,\ \forall \bar{p}' \in (0, p_{\max}]\right\}$$

and

$$\Lambda := \bigcup_{\bar{p} \in (0, p_{\max}]} \Lambda(\bar{p})\,.$$

Remark 5.5.1. *It follows from the definition of $\Lambda(\bar{p})$ that*

$$\inf_{\nu \in \mathcal{G}} \int_{\mathbb{R}_+ \times \widetilde{U}} \bigl[c(x) + \lambda h(\tilde{u})\bigr]\, \nu(dx, d\tilde{u}) = J^*(\bar{p}) + \lambda \bar{p}\,,$$

for all $\lambda \in \Lambda(\bar{p})$. Also, it is rather straightforward to show that $\Lambda = [0, \bar{\lambda})$ for some $\bar{\lambda} \in \mathbb{R}_+ \cup \{\infty\}$.

Lemma 5.5.3. *Let Assumption 5.4.1 hold. Then, for all $\bar{p} \in (0, p_{max}]$ and $\lambda \in \Lambda$,*

$$\liminf_{x \to \infty} \inf_{\tilde{u} \in \widetilde{U}} L(x, \tilde{u}, \bar{p}, \lambda) > \tilde{J}^*(\bar{p}, \lambda)\,. \tag{5.22}$$

Proof. If c is asymptotically unbounded, (5.22) always follows. Otherwise, fix $\bar{p} \in (0, p_{\max}]$ and $\lambda \in \Lambda$. Let $\bar{p}' \in (0, p_{\max}]$ be such that $\lambda \in \Lambda(\bar{p}')$. By convexity

$$J^*\bigl(\tfrac{\bar{p}'}{2}\bigr) \geq J^*(\bar{p}') + \lambda \tfrac{\bar{p}'}{2}\,.$$

Thus, using Lemma 5.5.2, we obtain

$$J^*(\bar{p}') + \lambda \bar{p}' < c_\infty\,. \tag{5.23}$$

Hence, by (5.21) and (5.23),

$$\begin{aligned}
\liminf_{x \to \infty} \inf_{\tilde{u} \in \widetilde{U}} L(x, \tilde{u}, \bar{p}, \lambda) + \lambda \bar{p} &= \liminf_{x \to \infty} c(x) \\
&> J^*(\bar{p}') + \lambda \bar{p}' \\
&= \tilde{J}^*(\bar{p}, \lambda) + \lambda \bar{p}\,,
\end{aligned}$$

and the proof is complete. ∎

5.5.1 The structure of the optimal policy

Using the theory in [15, Chapter IV.3], we can characterize optimality via the HJB equation. This is summarized as follows:

Theorem 5.5.4. *Let Assumption 5.4.1 hold. Fix $\bar{p} \in (0, p_{max}]$ and $\lambda_{\bar{p}} \in \Lambda(\bar{p})$. Then there exists a unique solution pair (V, β), with $V \in \mathcal{C}^2(\mathbb{R}_+, \mathbb{R})$ and $\beta \in \mathbb{R}$, to the HJB*

$$\min_{\tilde{u} \in \widetilde{U}} \left[\mathcal{L}^{\tilde{u}} V(x) + L(x, \tilde{u}, \bar{p}, \lambda_{\bar{p}})\right] = \beta\,, \tag{5.24a}$$

subject to the boundary condition

$$\frac{dV}{dx}(0) = 0\,, \tag{5.24b}$$

and also satisfying

(a) $V(0) = 0$

(b) $\inf_{x \in \mathbb{R}_+} V(x) > -\infty$

(c) $\beta \leq \tilde{J}^*(\bar{p}, \lambda_{\bar{p}})$.

Moreover, if v^* is a measurable selector of the minimizer in (5.24a)*, then $v^* \in \mathfrak{U}_{se} \subset \mathfrak{U}_{ss}$, and v^* is an optimal policy for* (5.20)*, or equivalently, for* (5.15)*. Also, $\beta = \tilde{J}^*(\bar{p}, \lambda_{\bar{p}}) = J^*(\bar{p})$ (the second equality follows by Theorem 5.5.1).*

Following [15, Chapter IV.1] we can show that the stationary policy v^* in Theorem 5.5.4 is optimal among all admissible controls $\mathfrak{U}$, and hence is a minimizer for (5.6). This is done as follows: For a control $v \in \mathfrak{U}$ define the process $\{\varphi^v_t\,,\ t \geq 0\}$ of empirical measures as a $\mathcal{P}(\mathbb{R}_+ \times \widetilde{U})$-valued process satisfying, for all $g \in \mathcal{C}_b(\mathbb{R}_+ \times \widetilde{U})$,

$$\varphi^v_t(A, B) = \frac{1}{t} \int_0^t I_A\big(x(s)\big) \eta_v\big(x(s), B\big)\, ds\,.$$

Suppose that $v \in \mathfrak{U}$ is such that, for $\bar{p} \in (0, p_{\max}]$,

$$\limsup_{t \to \infty}\ \frac{1}{t} \int_0^t h\big(v(s)\big)\, ds \leq \bar{p}\,, \quad \text{a.s.} \tag{5.25}$$

Following the approach in [15, Chapter IV.1], utilizing the near-monotone property asserted in Lemma 5.5.3 and the characterization of $\mathcal{G}$ in (5.12), we first deduce that any subsequence $\{t_n\}$, $t_n \to \infty$, contains a further subsequence $\{t'_n\}$ along which $\varphi^v_{t'_n}$ converges weakly, as $n \to \infty$, to some $\nu \in \mathcal{G}$. Thus

$$\begin{aligned}\limsup_{t\to\infty} \frac{1}{t}\int_0^t L\bigl(x(s), v(s), \bar{p}, \lambda_{\bar{p}}\bigr)\, ds \\ &\ge \int_{\mathbb{R}_+\times\tilde{U}} L\bigl(x, \tilde{u}, \bar{p}, \lambda_{\bar{p}}\bigr)\,\nu(dx, d\tilde{u}) \\ &\ge \tilde{J}^*(\bar{p}, \lambda_{\bar{p}})\,, \quad \text{a.s.}\end{aligned} \tag{5.26}$$

Then, (5.25)–(5.26) imply that under the policy v

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t c\bigl(x(s)\bigr)\, ds \ge \tilde{J}^*(\bar{p}, \lambda_{\bar{p}}) = J^*(\bar{p})\,, \quad \text{a.s.} \tag{5.27}$$

Optimality of $v^* \in \mathfrak{U}_{\mathrm{se}}$ then follows by (5.25) and (5.27), and we have the following theorem.

Theorem 5.5.5. *Under Assumption 5.4.1, for any $\bar{p} \in (0, p_{max}]$, there exists $v^* \in \mathfrak{U}_{se}$ which attains the minimum in* (5.6) *over all controls in $\mathfrak{U}$.*

If $\Lambda(\bar{p})$ and $J^*(\bar{p})$ were known, then one could solve (5.24) and derive the optimal policy. Since this is not the case, we embark on a different approach. We write (5.24) as

$$\min_{\tilde{u}\in\widetilde{U}} \bigl[\mathcal{L}^{\tilde{u}}V(x) + c(x) + \lambda_{\bar{p}} h(\tilde{u})\bigr] = \beta + \lambda_{\bar{p}}\,\bar{p}\,. \tag{5.28}$$

By Theorem 5.5.4, $J^*(\bar{p})$ is the smallest value of β for which there exists a solution pair (V, β) to (5.24), satisfying (b). This yields the following corollary:

Corollary 5.5.6. *Let Assumption 5.4.1 hold. For $\lambda \in \Lambda$, consider the HJB equation*

$$\min_{\tilde{u}\in\widetilde{U}} \bigl[\mathcal{L}^{\tilde{u}}V(x) + c(x) + \lambda h(\tilde{u})\bigr] = \varrho\,, \tag{5.29a}$$

subject to the boundary condition

$$\frac{dV}{dx}(0) = 0\,, \tag{5.29b}$$

and define

$$\mathcal{Q}_\lambda := \Big\{(V,\varrho) \text{ solves } (5.29) \text{ and } \inf_{x\in\mathbb{R}_+} V(x) > -\infty\Big\} \tag{5.30a}$$

$$\varrho_\lambda := \min\big\{\varrho : (V,\varrho)\in\mathcal{Q}_\lambda\big\}\,. \tag{5.30b}$$

Then

$$\varrho_\lambda = \min_{v\in\mathfrak{U}_{ss}} \int_{\mathbb{R}_+\times\tilde{U}} \big[c(x) + \lambda h(\tilde{u})\big]\,\nu_v(dx,d\tilde{u})\,. \tag{5.31}$$

Furthermore, if $\bar{p}$ is a point in $(0, p_{max}]$ such that $\lambda \in \Lambda(\bar{p})$, then $\varrho_\lambda = J^*(\bar{p}) + \lambda\bar{p}$, and if v^*_λ is a measurable selector of the minimizer in (5.29a) with $\varrho = \varrho_\lambda$, then v^*_λ is a stationary optimal policy for (5.20).

The minimizer in (5.29a) satisfies

$$\min_{\tilde{u}\in\tilde{U}} \Big[-b(\tilde{u})\frac{dV}{dx} + \lambda h(\tilde{u})\Big] = \min_{\tilde{u}\in\tilde{U}} \sum_j (\lambda - \gamma_j\frac{dV}{dx})\pi_j\tilde{u}_j\,.$$

Thus the optimal control v^*_λ takes the following simple form: for $i = 1,\dots,N$ and $x\in\mathbb{R}_+$,

$$(v^*_\lambda)_i(x) = \begin{cases} 0\,, & \text{if } \gamma_i\frac{dV}{dx}(x) < \lambda \\ p_{\max}\,, & \text{if } \gamma_i\frac{dV}{dx}(x) \ge \lambda\,. \end{cases} \tag{5.32}$$

Thus, provided $\frac{dV}{dx}$ is monotone, the optimal control v^*_λ is of multi-threshold type, i.e., for each channel state j there is a queue-threshold $\hat{x}_j$, such that at any time t, the optimal policy transmits at peak power $p_{\max}$ over channel state j, if the queue length $x(t) > \hat{x}_j$, and does not transmit otherwise.

Further, from Remark 5.3.2, it follows that if the equilibrium power $\{P_0(j)\}$ is allocated according to channel-state dependent water-filling with strictly positive equilibrium power allocations for each channel state, the multi-threshold policy collapses to a *single-threshold* policy (since $\gamma_i = \gamma_j$, for all i, j). In other words, there is a state-independent queue-threshold $\hat{x}$, such that at any time t, the optimal policy transmits at peak power $p_{\max}$, if the queue length $x(t) > \hat{x}$, and does not transmit otherwise.

The following lemma asserts the monotonicity of $\frac{dV}{dx}$, under the additional assumption that c is non-decreasing.

Lemma 5.5.7. *Suppose c satisfies Assumption 5.4.1, and is non-decreasing on $[0,\infty)$. Then every $(V,\varrho) \in \mathcal{Q}_\lambda$ satisfies*

(a) *$\frac{dV}{dx}$ is non-decreasing;*

(b) *If c is unbounded, then $\frac{dV}{dx}$ is unbounded.*

Proof. Equation (5.29a) takes the form

$$\frac{\sigma^2}{2}\frac{d^2V}{dx^2}(x) = \sum_j \pi_j p_{\max}\Big[\gamma_j \frac{dV}{dx}(x) - \lambda\Big]^+ + \varrho - c(x)\,, \tag{5.33}$$

where the initial condition is given by (5.29b). Since c is non-decreasing, then by (5.31), $\varrho > c(0)$. Suppose that for some $x' \in \mathbb{R}_+$, $\frac{d^2V}{dx^2}(x') = -\varepsilon < 0$. Let $x'' = \inf\big\{x > x' : \frac{d^2V}{dx^2}(x) \geq 0\big\}$. Since by Theorem 5.5.4 $\frac{d^2V}{dx^2}$ is continuous, it must hold $x'' > x'$. Suppose $x'' < \infty$. Since $\frac{d^2V}{dx^2} < 0$ on $[x', x'')$ and $\varrho - c(x)$ is non-increasing, (5.33) implies that $\frac{d^2V}{dx^2}(x'') \leq \frac{d^2V}{dx^2}(x') < 0$. Thus we are led to a contradiction, and it follows that $\frac{d^2V}{dx^2}(x) \leq -\varepsilon < 0$, for all $x \in [x', \infty)$, implying that V is not bounded below. It is clear from (5.33) that since $\frac{d^2V}{dx^2} \geq 0$, then $\frac{d^2V}{dx^2}(x) \to \infty$, as $x \to \infty$, provided c is not bounded. □

The proof of Lemma 5.5.7 shows that if (V,ϱ) solves (5.29), then V is bounded below, if and only if $\frac{d^2V}{dx^2}(x) \geq 0$, for all $x \in \mathbb{R}_+$. Thus $\mathcal{Q}_\lambda$ defined in (5.30a), has an alternate characterization given in the following corollary.

Corollary 5.5.8. *Suppose c satisfies Assumption 5.4.1, and is non-decreasing on $[0,\infty)$. Then, for all $\lambda \in \Lambda$,*

$$\mathcal{Q}_\lambda = \Big\{(V,\varrho) \ \textit{solves}\ (5.29)\ \textit{and}\ \tfrac{d^2V}{dx^2} \geq 0\,,\ \textit{on}\ \mathbb{R}_+\Big\}\,.$$

Comparing (5.29) and (5.28), a classical application of Lagrange duality (see [67, pg. 224, Thm. 1]) yields the following:

Lemma 5.5.9. *If c satisfies Assumption 5.4.1, and is non-decreasing on* $[0,\infty)$, *then, for any $\bar{p} \in (0, p_{max}]$ and $\lambda_{\bar{p}} \in \Lambda(\bar{p})$, we have:*

$$\varrho_{\lambda_{\bar{p}}} - \lambda_{\bar{p}}\,\bar{p} = \max_{\lambda \geq 0} \left\{ \varrho_\lambda - \lambda \bar{p} \right\} = J^*(\bar{p})\,. \tag{5.34}$$

Moreover, if λ_0 attains the maximum in $\lambda \mapsto \varrho_\lambda - \lambda\bar{p}$ then $\varrho_{\lambda_0} = J^*(\bar{p}) + \lambda_0 \bar{p}$, which implies that $\lambda_0 \in \Lambda(\bar{p})$.

Remark 5.5.2. *Lemma 5.5.9 furnishes a method for solving* (5.15). *This can be done as follows: With λ viewed as a parameter, we first solve for ϱ_λ which is defined in* (5.30b). *Then, given $\bar{p}$, we obtain the corresponding value of the Lagrange multiplier via the maximization in* (5.34). *The optimal control can then be evaluated using* (5.32), *with $\lambda = \lambda_{\bar{p}}$. Section 5.6.1 contains an example demonstrating this method.*

5.6 Solution of the HJB equation

In this section we present an analytical solution of the HJB equation (5.29). We deal only with the case where the cost function c is non-decreasing and asymptotically unbounded. However, the only reason for doing so is in the interest of simplicity and clarity. If c is bounded the optimal policy may have less than N threshold points, but other than the need to introduce some extra notation, the solution we outline below for unbounded c, holds virtually unchanged for the bounded case. Also, without loss of generality, we assume that $\gamma_1 > \cdots > \gamma_N > 0$.

We parameterize the policies in (5.32) by a collection of points $\{\hat{x}_1, \ldots, \hat{x}_N\}$ in $\mathbb{R}_+$. In other words, if V is the solution (5.33), then $\hat{x}_i$ is the least positive number such that $\frac{dV}{dx}(\hat{x}_i) \geq \gamma_i^{-1}$. Thus, if we define

$$\mathcal{X}^N := \left\{ \hat{x} = (\hat{x}_1, \ldots, \hat{x}_N) \in \mathbb{R}_+^N : \hat{x}_1 < \cdots < \hat{x}_N \right\},$$

then for each $\hat{x} \in \mathcal{X}^N$, there corresponds a multi-threshold policy $v_{\hat{x}}$ of the form

$$(v_{\hat{x}})_i(x) = \begin{cases} p_{\max}\,, & \text{if} \quad x \geq \hat{x}_i \\ 0\,, & \text{otherwise.} \end{cases} \qquad 1 \leq i \leq N\,. \tag{5.35}$$

To facilitate expressing the solution of (5.33), we need to introduce some new notation. For $i = 1, \ldots, N$, define

$$\tilde{\pi}_i := \sum_{j=1}^{i} \pi_i \,, \quad \tilde{\gamma}_i := \sum_{j=1}^{i} \pi_i \gamma_i \,, \quad \mathit{\Gamma}_i := \frac{\tilde{\gamma}_i}{\gamma_i} - \tilde{\pi}_i \,.$$

Note that from (5.14), we obtain the identity

$$\alpha_i = \frac{2p_{\max}}{\sigma^2} \tilde{\gamma}_i \,, \quad i = 1, \ldots, N \,.$$

For $x, z \in \mathbb{R}_+$, with $z \le x$, we define the functions

$$F_0(\varrho, x) := \varrho x - \int_0^x c(y)\, dy \,,$$

and for $i = 1, \ldots, N$,

$$F_i(\varrho, x, z) := [\varrho + \lambda p_{\max} \mathit{\Gamma}_i]\big(1 - e^{\alpha_i (z-x)}\big) - \alpha_i \int_z^x e^{\alpha_i (z-y)} c(y)\, dy \,,$$
$$G_i(\varrho, x, z) := \varrho + \lambda p_{\max} \mathit{\Gamma}_i - \alpha_i \int_z^x e^{\alpha_i (z-y)} c(y)\, dy - e^{\alpha_i (z-x)} c(x) \,.$$

Using the convention $\hat{x}_{N+1} \equiv \infty$, we write the solution of (5.33) as

$$\frac{dV}{dx}(x) = \frac{2}{\sigma^2} F_0(\varrho, x) \,, \quad 0 \le x < \hat{x}_1 \,, \tag{5.36a}$$

and for $x \in [\hat{x}_i, \hat{x}_{i+1})$, $i = 1, \ldots, N$,

$$\frac{dV}{dx}(x) = \frac{2}{\sigma^2 \alpha_i} e^{\alpha_i (x - \hat{x}_i)} F_i(\varrho, x, \hat{x}_i) + \frac{\lambda}{\gamma_i} \,. \tag{5.36b}$$

In addition, the following boundary conditions are satisfied

$$F_0(\varrho, \hat{x}_1) - \frac{\lambda \sigma^2}{2\gamma_1} = 0 \,, \tag{5.37a}$$

and for $i = 1, \dots, N-1$,

$$F_i(\varrho, \hat{x}_{i+1}, \hat{x}_i) = \lambda p_{\max} \tilde{\gamma}_i e^{\alpha_i(\hat{x}_i - \hat{x}_{i+1})} \Big(\frac{1}{\gamma_{i+1}} - \frac{1}{\gamma_i} \Big) . \tag{5.37b}$$

Also, for $i = 1, \dots, N$, we have

$$\frac{d^2 V}{dx^2}(x) = \frac{2}{\sigma^2} e^{\alpha_i (x - \hat{x}_i)} G_i(\varrho, x, \hat{x}_i) , \quad x \in (\hat{x}_i, \hat{x}_{i+1}) .$$

Since c is monotone, the map

$$x \mapsto \alpha_i \int_z^x e^{\alpha_i(z-y)} c(y)\, dy + e^{\alpha_i(z-x)} c(x) \tag{5.38}$$

is non-decreasing. Moreover, using the fact that c is either asymptotically unbounded (or strictly monotone increasing, when bounded), an easy calculation yields

$$G_i(\varrho, x, z) > \lim_{x \to \infty} G_i(\varrho, x, z) . \tag{5.39}$$

Suppose $\hat{x} \in \mathcal{X}^N$, are the threshold points of a solution (V, ϱ) of (5.33). It follows from (5.39) that $\lim_{x\to\infty} G_N(\varrho, x, \hat{x}_N) \geq 0$ is a necessary and sufficient condition for $\frac{d^2V}{dx^2}(x) \geq 0$, for all $x \in (\hat{x}_N, \infty)$. This condition translates to

$$\varrho + \lambda p_{\max} \mathit{\Gamma}_N - \alpha_N \int_{\hat{x}_N}^{\infty} e^{\alpha_N(\hat{x}_N - y)} c(y)\, dy \geq 0 . \tag{5.40}$$

The arguments in the proof of Lemma 5.5.7 actually show that (5.40) is sufficient for $\frac{d^2V}{dx^2}$ to be non-negative on $\mathbb{R}_+$. We sharpen this result by showing in Lemma 5.6.1 below that (5.40) implies that $\frac{d^2V}{dx^2}$ is strictly positive on $\mathbb{R}_+$.

Lemma 5.6.1. *Suppose $\hat{x} \in \mathcal{X}^N$ satisfies* (5.37). *If* (5.40) *holds, then $\varrho > c(\hat{x}_1)$ and $G_i(\varrho, x, \hat{x}_i) > 0$, for all $x \in [\hat{x}_i, \hat{x}_{i+1}]$, $i = 0, \dots, N-1$.*

Proof. We argue by contradiction. If $\varrho \leq c(\hat{x}_1)$, then $G_1(\varrho, \hat{x}_i, \hat{x}_i) \leq 0$, hence it is enough to assume that $G_i(\varrho, x, \hat{x}_i) \leq 0$, for some $x \in [\hat{x}_i, \hat{x}_{i+1}]$ and $i \in \{1, \dots, N-1\}$. Then, since (5.38) is non-decreasing,

$$G_i(\varrho, \hat{x}_{i+1}, \hat{x}_i) \leq 0 . \tag{5.41}$$

Therefore, since

$$F_i(\varrho, x, \hat{x}_i) = G_i(\varrho, x, \hat{x}_i) + e^{\alpha_i(\hat{x}_i - x)} c(x) - [\varrho + \lambda p_{\max} \mathit{\Gamma}_i] e^{\alpha_i(\hat{x}_i - x)}, \qquad (5.42)$$

combining (5.37b) and (5.41)–(5.42), we obtain

$$c(\hat{x}_{i+1}) - \varrho - \lambda p_{\max} \mathit{\Gamma}_i \geq \lambda \tilde{\gamma}_i \Big(\frac{1}{\gamma_{i+1}} - \frac{1}{\gamma_i} \Big),$$

which simplifies to

$$c(\hat{x}_{i+1}) - \varrho + \lambda p_{\max} \tilde{\pi}_i \geq \lambda p_{\max} \frac{\tilde{\gamma}_i}{\gamma_{i+1}}. \qquad (5.43)$$

Since

$$\frac{\tilde{\gamma}_i}{\gamma_{i+1}} - \tilde{\pi}_i = \frac{\tilde{\gamma}_{i+1}}{\gamma_{i+1}} - \tilde{\pi}_{i+1} = \mathit{\Gamma}_{i+1},$$

(5.43) yields

$$\varrho + \lambda p_{\max} \mathit{\Gamma}_{i+1} \leq c(\hat{x}_{i+1}). \qquad (5.44)$$

Using the monotonicity of $x \mapsto G_{i+1}(\varrho, x, \hat{x}_{i+1})$ together with (5.44), we get

$$G_{i+1}(\varrho, x, \hat{x}_{i+1}) \leq 0,$$

for all $x \in [\hat{x}_{i+1}, \hat{x}_{i+2}]$, and iterating this argument, we conclude that

$$G_N(\varrho, x, \hat{x}_N) \leq 0,$$

for all $x \in (\hat{x}_N, \infty)$, thus contradicting (5.40). □

Combining Corollary 5.5.8 with Lemma 5.6.1, yields the following.

Corollary 5.6.2. *Suppose* (V, ϱ) *satisfies* (5.36)–(5.37), *for some* $\hat{x} \in \mathcal{X}^N$ *and* $\lambda \in \mathit{\Lambda}$. *Then* $(V, \varrho) \in \mathcal{Q}_\lambda$, *if and only if* (5.40) *holds.*

For $\lambda \in \mathit{\Lambda}$, define

$$\mathcal{R}_\lambda := \{\varrho \in \mathbb{R}_+ : (V, \varrho) \in \mathcal{Q}_\lambda\}.$$

For each $\varrho \in \mathcal{R}_\lambda$, equations (5.37) define a map $\varrho \mapsto \hat{x}$, which we denote by $\hat{x}(\varrho)$.

Lemma 5.6.3. *Let $\lambda \in \Lambda$ and suppose $\varrho_0 \in \mathcal{R}_\lambda$. With ϱ_λ as defined in* (5.30b)*, and denoting the left-hand side of* (5.40) *by $G_N(\varrho, \infty, \hat{x}_N)$, the following hold:*

(a) *If $\varrho' > \varrho_0$, then $\varrho' \in \mathcal{R}_\lambda$ and $G_N\big(\varrho', \infty, \hat{x}(\varrho')\big) > 0$.*

(b) *If $G_N\big(\varrho_0, \infty, \hat{x}(\varrho_0)\big) > 0$, then $\varrho_0 > \varrho_\lambda$.*

(c) *$\mathcal{R}_\lambda = [\varrho_\lambda, \infty)$, and ϱ_λ is the only point in $\mathcal{R}_\lambda$ which satisfies $G_N\big(\varrho_\lambda, \infty, \hat{x}(\varrho_\lambda)\big) = 0$.*

Proof. Part (a) follows easily from (5.33). Denoting by V_0 and V' the solutions of (5.33) corresponding to ϱ_0 and ϱ', respectively, a standard argument shows that

$$\frac{d^2(V' - V_0)}{dx^2}(x) \geq \varrho' - \varrho_0 > 0\,, \quad \forall x \in \mathbb{R}_+\,,$$

implying

$$\frac{dV'}{dx}(x) \geq \frac{dV_0}{dx}(x)\,, \quad \forall x \in \mathbb{R}_+\,. \tag{5.45}$$

Hence, since by the definition of $\mathcal{Q}_\lambda$, V_0 is bounded below, the same holds for V', in turn implying that $(V', \varrho') \in \mathcal{Q}_\lambda$. By (5.45), $\hat{x}(\varrho') \leq \hat{x}(\varrho_0)$, and since $\hat{x}_N \mapsto G_N(\varrho, \infty, \hat{x}_N)$ is non-increasing and $\varrho' > \varrho_0$, we obtain $G_N\big(\varrho', \infty, \hat{x}(\varrho')\big) > 0$.

Concerning (b), we write (5.37) in the form $\tilde{F}(\varrho, \hat{x}) = 0$, with $\tilde{F} : \mathbb{R}_+^{N+1} \to \mathbb{R}_+^N$. The map $\tilde{F}$ is continuously differentiable and as a result of Lemma 5.6.1 its Jacobian $D_{\hat{x}}F$ with respect to $\hat{x}$ has full rank at $\big(\varrho_0, \hat{x}(\varrho_0)\big)$. By the Implicit Function Theorem, there exists an open neighborhood $W(\varrho_0)$ and a continuous map $\hat{x} : W(\varrho_0) \to \mathbb{R}_+$, such that $\tilde{F}\big(\varrho, \hat{x}(\varrho)\big) = 0$, for all $\varrho \in W(\varrho_0)$. Using the continuity of G_N, we may restrict $W(\varrho_0)$ further so that $G_N\big(\varrho, \infty, \hat{x}(\varrho)\big) > 0$, for all $\varrho \in W(\varrho_0)$. Hence $W(\varrho_0) \subset \mathcal{R}_\lambda$, implying that $\varrho_0 > \varrho_\lambda$.

Part (c) follows directly from (a) and (b). ∎

Combining Corollary 5.5.6 and Lemma 5.6.1, we obtain the following characterization of the solution to the HJB equation (5.29).

Theorem 5.6.4. *Let c be non-decreasing and asymptotically unbounded. Then, the threshold points $(\hat{x}_1, \ldots, \hat{x}_N) \in \mathcal{X}^N$ of the stationary optimal policy in* (5.35) *and the optimal value $\varrho_\lambda > 0$, are the (unique) solution of the set of $N+1$ algebraic equations which is comprised of the equations in* (5.37) *and $G_N\big(\varrho_\lambda, \infty, \hat{x}(\varrho_\lambda)\big) = 0$.*

5.6.1 Example: minimizing the mean delay

We specialize the optimization problem to the case $c(x) = x$, which corresponds to minimizing the mean delay.

First consider the case $N = 1$, letting $\alpha \equiv \alpha_1$ and $\hat{x} \equiv \hat{x}_1$. Solving (5.29) we obtain

$$\frac{dV}{dx}(x) = \frac{2\varrho}{\sigma^2}x - \frac{x^2}{\sigma^2}\,, \quad x \leq \hat{x}\,,$$

with

$$\hat{x} = \varrho - \sqrt{\varrho^2 - \tfrac{\lambda\sigma^2}{\gamma}}\,. \tag{5.46}$$

Also, for $x \geq \hat{x}$,

$$\frac{dV}{dx}(x) = \frac{2e^{\alpha(x-\hat{x})}}{\sigma^2\alpha}\Big(\varrho - \hat{x} - \frac{1}{\alpha}\Big) + \frac{2}{\sigma^2\alpha}\Big(\varrho - \lambda p_{\max} + x + \frac{1}{\alpha}\Big)\,.$$

Therefore, for $x > \hat{x}$,

$$\frac{d^2V}{dx^2}(x) = \frac{2}{\sigma^2}\Big(\varrho - \hat{x} - \frac{1}{\alpha}\Big)e^{\alpha(x-\hat{x})} + \frac{2}{\sigma^2\alpha}\,. \tag{5.47}$$

It follows from (5.47) that

$$\varrho_\lambda = \hat{x} + \frac{1}{\alpha}\,. \tag{5.48}$$

By (5.46) and (5.48),

$$\varrho_\lambda = \sqrt{\tfrac{1}{\alpha^2} + \tfrac{\lambda\sigma^2}{\gamma}}\,. \tag{5.49}$$

Let $\bar{p} \in (0, p_{\max}]$ be given. Applying Lemma 5.5.9, we obtain from (5.49)

$$\lambda_{\bar{p}} = \frac{p_{\max}}{2\alpha\bar{p}^2} - \frac{1}{2\alpha p_{\max}}\,.$$

and

$$J^*(\bar{p}) = \frac{1}{2\alpha}\Big(\frac{p_{\max}}{\bar{p}} + \frac{\bar{p}}{p_{\max}}\Big)\,.$$

Moreover, the threshold point of the optimal policy is given by

$$\hat{x} = \frac{1}{\alpha}\Big(\frac{p_{\max}}{\bar{p}} - 1\Big)\,. \tag{5.50}$$

Now consider the case $N = 2$. We obtain:

$$\frac{dV}{dx}(x) = \frac{2\varrho}{\sigma^2}x - \frac{x^2}{\sigma^2}\,, \quad x \le \hat{x}_1 \tag{5.51a}$$

$$\begin{aligned}\frac{dV}{dx}(x) &= \frac{2}{\sigma^2\alpha_1}\Big(\varrho - \hat{x}_1 - \frac{1}{\alpha_1}\Big)\big[e^{\alpha_1(x-\hat{x}_1)} - 1\big] \\ &\qquad + \frac{2(x-\hat{x}_1)}{\sigma^2\alpha_1} + \frac{\lambda}{\gamma_1}\,, \quad \hat{x}_1 \le x < \hat{x}_2\,,\end{aligned} \tag{5.51b}$$

and for $x \ge \hat{x}_2$,

$$\begin{aligned}\frac{dV}{dx}(x) &= \frac{2}{\sigma^2\alpha_2}\Big(\varrho - \hat{x}_2 - \frac{1}{\alpha_2} + \lambda p_{\max}\pi_1\frac{\gamma_1-\gamma_2}{\gamma_2}\Big) \\ &\qquad \times \big[e^{\alpha_2(x-\hat{x}_2)} - 1\big] + \frac{2(x-\hat{x}_2)}{\sigma^2\alpha_2} + \frac{\lambda}{\gamma_2}\,.\end{aligned} \tag{5.51c}$$

Since $\frac{dV}{dx}(\hat{x}_1) = \frac{\lambda}{\gamma_1}$, we obtain by (5.51a),

$$\hat{x}_1 = \varrho - \sqrt{\varrho^2 - \tfrac{\lambda\sigma^2}{\gamma_1}}\,. \tag{5.52}$$

By (5.51c), $\frac{d^2V}{dx^2}(x) \ge 0$, for all $x > \hat{x}_2$, if and only if

$$\varrho - \hat{x}_2 - \frac{1}{\alpha_2} + \lambda p_{\max}\pi_1\frac{\gamma_1-\gamma_2}{\gamma_2} \ge 0\,.$$

Also, since $\frac{dV}{dx}(\hat{x}_2) = \frac{\lambda}{\gamma_2}$, we obtain from (5.51b),

$$\Big(\varrho - \hat{x}_1 - \frac{1}{\alpha_1}\Big)\big[e^{\alpha_1(\hat{x}_2-\hat{x}_1)} - 1\big] + \hat{x}_2 - \hat{x}_1 = \frac{\sigma^2\lambda\alpha_1}{2}\Big(\frac{1}{\gamma_2} - \frac{1}{\gamma_1}\Big)\,.$$

We apply Theorem 5.6.4 to compute the optimal policy. Define $\hat{x}_1(\varrho)$ by (5.52) and

$$\hat{x}_2(\varrho) := \hat{x}_1(\varrho) + \sqrt{\varrho^2 - \tfrac{\lambda\sigma^2}{\gamma_1}} - \tfrac{1}{\alpha_2} + \lambda p_{\max}\pi_1\tfrac{\gamma_1-\gamma_2}{\gamma_2}\,.$$

Then ϱ_λ is the solution of

$$\Big(\sqrt{\varrho^2 - \tfrac{\lambda\sigma^2}{\gamma_1}} - \tfrac{1}{\alpha_1}\Big)e^{\alpha_2(\hat{x}_2(\varrho)-\hat{x}_1(\varrho))} + \Big(\tfrac{1}{\alpha_1} - \tfrac{1}{\alpha_2}\Big) = 0\,.$$

In Figure 5.4 we plot the optimal threshold points for a two state channel

$(N = 2)$ as a function of $\bar{p}$. The parameters are selected as $\pi = (0.5, 0.5)$, $\gamma = (2, 1)$, $\sigma = 1$ and $p_{\max} = 1$.

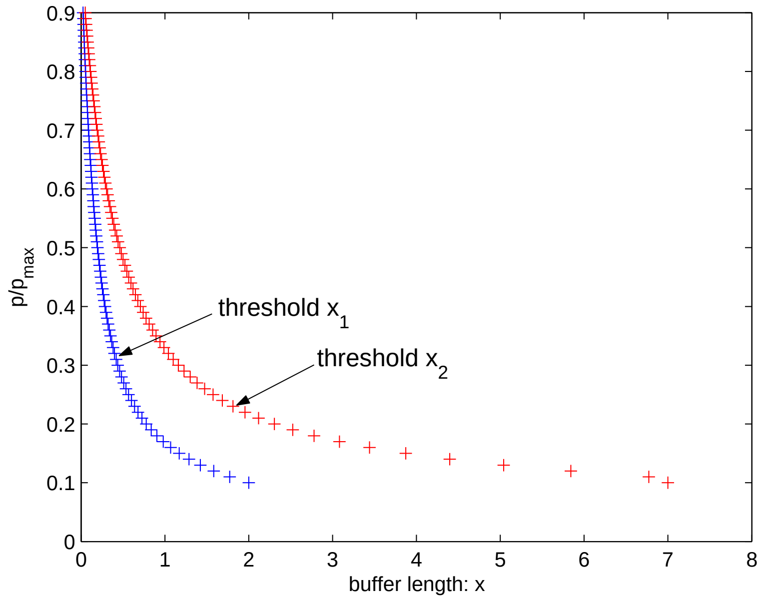


Figure 5.4: Optimal threshold points as a function of $\bar{p}$.

5.7 Numerical results

We have considered the optimal power allocation problem in a time-varying channel under the heavy-traffic approximation. In the heavy-traffic region, the queueing process is modeled as a controlled diffusion process. The policy which minimizes the delay subject to a long-term average power constraint is multi-threshold and can be computed by the procedure outlined in Theorem 5.6.4. In this section, we compare the performance of the optimal policy under the heavy-traffic approximation with the optimal policy for the original non-scaled system. The latter is computed numerically in [11].

In [11], under the Poisson assumption on the arrival process, the power allocation problem is formulated as a discrete-time Markov decision process (MDP) with the state variable (X, g), where X is the buffer state, g is the channel state, and the action $P(X, g)$ is the transmitting power. With $A(t)$ denoting the arrival process, the queueing process is described by

$$X(t) = \min \left\{ \max \left\{ X(t-1) + A(t) - D(t), 0 \right\}, L \right\},$$

where L is the buffer size, and the departure process $D(t)$ is controlled by the power allocation $P(X, g)$.

In our simulations, we consider the power allocation in a two-state Markov channel with stationary distribution $\pi = [0.8, 0.2]$ and corresponding channel gains $g = [0.9, 0.3]$. The arrival process is a Poisson process with expectation $\lambda^a = 5$, and the service rate r depends on the power allocation P according to $r(P, g) = 10 \ln\left(1 + \frac{Pg}{10}\right)$.

Importantly, we comment here that the threshold based policy *does not* necessarily need a Poisson assumption for the proof of asymptotic optimality. For any sequence of arrival processes which converges to a Wiener process in the heavy-traffic limit, the threshold-based policy is asymptotically optimal. However, we do not know what the optimal policy is in the non-asymptotic regime with general arrivals. Thus, in our simulations, we compare the threshold-based policy with the optimal policy (obtained in [11]) with Poisson arrivals.

The numerical computation of the optimal policy of MDP in [11] is facilitated by standard methods, such as policy iteration and value iteration [81]. The optimal policies under different power constraints, are simulated to yield different average queue length drawn as the solid line in Figure 5.5. Note that the optimal policy under the heavy-traffic approximation is a single-threshold one. The optimal threshold as a function of the average power constraint can be obtained by (5.50). By using the threshold policies corresponding to different power constraints, a simulated power - queue length curve is plotted in Figure 5.5 with cross marks. The dotted line at the bottom in Figure 5.5 is the minimum power ($P_{\min} = 7.7$) required for the arrival rate to match the service rate (see (5.2)). By the affine relation between mean delay and mean queue length through Little's law with the constant of proportionality being the arrival rate, Figure 5.5 can be interpreted as a delay-power tradeoff curve. As can be seen in Figure 5.5, the two power-delay trade-off curves are very close, and they get even closer as the average queue length approaches $+\infty$, or equivalently, as the average power approaches $P_{\min}$, i.e., the heavy-traffic regime.

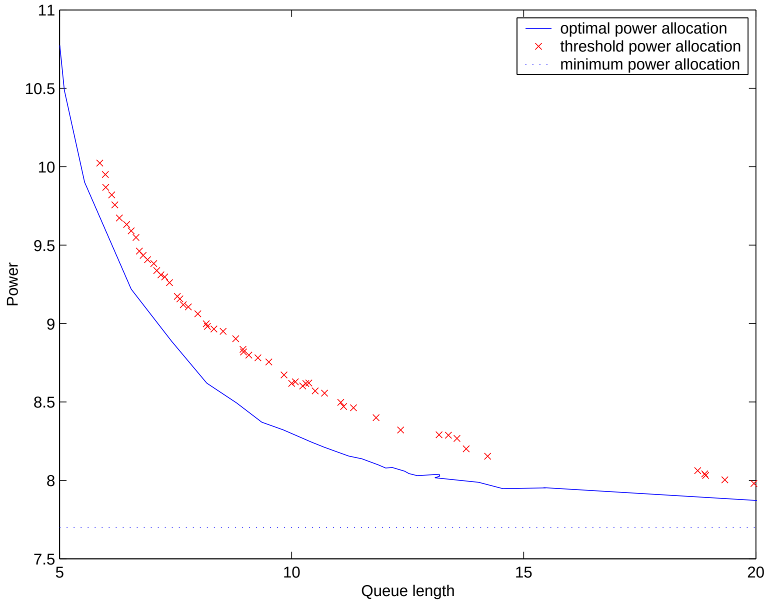


Figure 5.5: Power-delay trade-off curve comparison.

In terms of computational effort, in order to obtain the optimal policy of the discrete-time Markov decision process in [11] by value iteration or policy iteration, the complexity grows in proportion to the buffer size L, the number of channel states, the number of power levels, and the iteration steps needed, whereas the algorithm in Theorem 5.6.4 has complexity proportional to the number of channel states. With limited performance degradation, the multi-threshold policy has much simpler structure and lower computational complexity than the optimal control, and this makes it very promising for practical deployment.

Finally, the approach we have taken in this chapter for the resource allocation in a time-varying environment under heavy traffic approximation can be best summarized by Figure 5.6. For a general problem of minimizing average queueing cost under resource constraints, one can decouple it into two sub-problems: one is "basic allocation", basically a resource planning problem that minimizes the resource consumption subject to the service rates are no less than the arrival rates; another one is "extra allocation", which can be formulated as an optimal control problem of reflected diffusions under heavy traffic approximation. This essentially comes from a time-scale decomposition and yields an approximation for the true optimal solution.

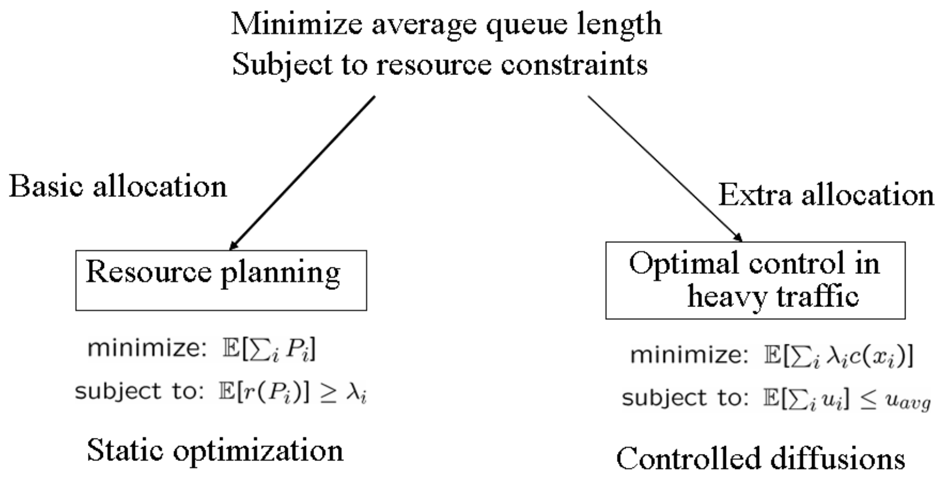


Figure 5.6: Illustration of the approach taken for resource allocation under heavy traffic approximation

Chapter 6

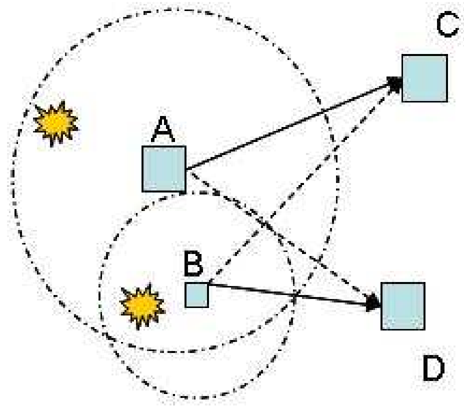
Cognitive Transmission: Information Theoretic View on Node Cooperation in Interference Channels

6.1 Introduction

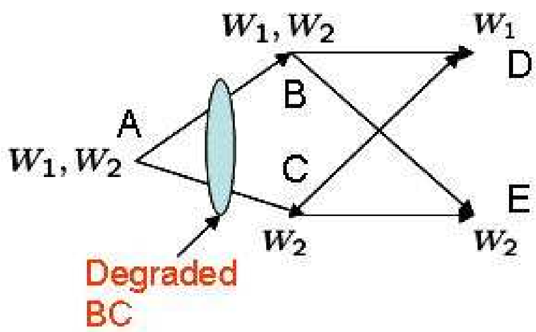
The interference channel (IFC) is a basic building block of most wireless networks, and is thus considered a fundamentally important channel from both theoretical and practical perspectives. However, the capacity region of this channel remains an open problem, with only some special cases being solved to date. Our goal is to investigate the capacity of a class of IFCs where one transmitter has full knowledge of the other transmitter's message in both the discrete memoryless and Gaussian cases. We term this class of channels "interference channels with degraded message sets".

IFCs with degraded message sets arise in many fairly important scenarios in wireless networks. The first is the cognitive radio channel introduced in [30] [1]. In this model, a cognitive transmitter gains full knowledge of another "dumb" transmitter's message. Each transmitter has a separate receiver associated with it. In this setting,

[1] Concurrent work in [52] also analyzes this channel.



(a) Data collection through IFC



(b) Cascaded with degraded broadcast channel

Figure 6.1: Two applications of IFCs with degraded message sets

the cognitive transmitter-receiver pair exploit the cognitive transmitter's knowledge of both messages to improve overall system performance. The second motivation for this problem lies in sensor networks as illustrated in Figure 6.1(a). In this setting, Sensor A has better sensing capability than Sensor B and thus can detect both events, while Sensor B can only detect one of them. In this setting, we assume that each sensor is aware of the capabilities of the other sensor and that the collected data needs to be sent to different receivers. Another motivation for this problem arises from the sensor network shown in Figure 6.1(b), in which the nodes B and C are constrained to fully decode the messages received from A. If the channel from A to B and C is a degraded broadcast channel, then the resulting channel from B and C to the receivers resembles an IFC with degraded message sets.

A central aspect of this channel is that the degraded structure of the messages allows the two transmitters to *cooperate*. Cooperation among transmitters to improve achievable rates has received considerable attention [49, 51, 76], particularly in an IFC setting [68] [69]. Our goal is twofold: first, to develop a cooperative encoding scheme that, for a class of discrete memoryless channels, achieves the capacity of the channel; second, to characterize the capacity region of a class of Gaussian IFC with degraded message sets with weak interference.

We begin by discussing the discrete memoryless case, where we find inner and outer bounds on the capacity of a class of IFCs with degraded message sets. Further, we determine conditions under which these two bounds meet. Next, we proceed to analyze the Gaussian IFC with degraded message sets. For a class of these channels where the inherent interference structure is "weak", we determine the capacity region.

The rest of the chapter is organized as follows. In Section 6.2, basic notations and definitions are introduced. The main results are presented in Section 6.3, including the capacity region of a class of discrete memoryless IFCs and Gaussian weak IFCs with degraded message sets. Numerical results of the Gaussian case are shown in Section 6.4.

6.2 Notations and Preliminaries

6.2.1 Channel Model and Definitions

We adopt the following notational conventions. Random variables (RVs) are denoted by capital letters, and their realizations by the respective lower case letters. X_m^n denotes the random vector $(X_m, \ldots, X_n)$, and X^n denotes the random vector $(X_1, \ldots, X_n)$. We use the notation

$$X \Rightarrow Y \Rightarrow Z$$

to denote that X and Z are conditionally independent given Y.

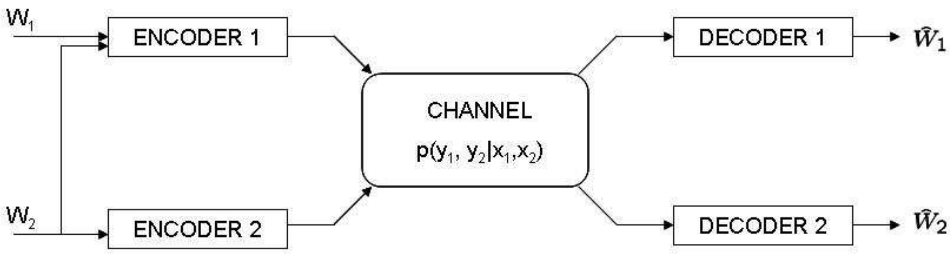


Figure 6.2: The model of an IFC with degraded message sets

A two-user IFC $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, p)$ is a channel with two input alphabets $\mathcal{X}_1$, $\mathcal{X}_2$, output alphabets $\mathcal{Y}_1$, $\mathcal{Y}_2$, and transition probability $p(y_1, y_2 \mid x_1, x_2)$. It is assumed that the channel is memoryless, namely

$$p(y_1^n, y_2^n \mid x_1^n, x_2^n) = \prod_{i=1}^{n} p(y_{1,i}, y_{2,i} \mid x_{1,i}, x_{2,i})\,.$$

Transmitter t sends a message W_t having M_t bits, to Receiver t, in n channel usages at rate $R_t = {}^{M_t}\!/_{n}$ bits per usage. A $(R_1, R_2, n, P_{e,1}, P_{e,2})$ code is defined as any code achieving the rate pair (R_1, R_2) with block size n and decoding error probability $P_{e,t}^{(n)}$, $t = 1, 2$. The *capacity region* $\mathcal{C}_{\text{IFC}}$ is the closure of the set of rate pairs (R_1, R_2), for which the receivers can decode their messages with error probability $P_{e,t}^{(n)} \to 0$ for $t = 1, 2$, as the block size $n \to \infty$.

In the classic IFC framework described above, each transmitter has its own message set $\mathcal{M}_t = \{W_t\}$, where $W_t \in \{1, 2, \dots, 2^{nR_t}\}$ denotes the private message to Receiver t. For an IFC with common information, a model proposed recently by Maric, Yates and Kramer in [68], each transmitter has not only its own private message W_t, but also the common message W_0 shared by all the transmitters. Thus the message set for Transmitter t is $\mathcal{M}_t = \{W_0, W_t\}$.

In our work, we consider IFCs *with degraded message sets* (IFC-DMS). For a two-user IFC-DMS, the message set of one transmitter is a strict subset of the other. For example, Figure 6.2 corresponds to the message sets,

$$\{W_2\} = \mathcal{M}_2 \subset \mathcal{M}_1 = \{W_1, W_2\}, \tag{6.1}$$

for which the capacity region is denoted as $\mathcal{C}_{\text{IFC}}^{T_1}$ to indicate that Transmitter 1 knows both messages. Note that Receiver t desires to decode Message W_t alone. Also note that $\mathcal{M}_t$ is the set of messages available to Transmitter t, which is in general a superset of W_t. Similarly, $\mathcal{C}_{\text{IFC}}^{T_2}$ denotes the capacity region of an IFC with Transmitter 2 knowing both messages, namely

$$\{W_1\} = \mathcal{M}_1 \subset \mathcal{M}_2 = \{W_1, W_2\}.$$

In recent works in the literature (e.g., [69]), an IFC-DMS has also been referred to as an *IFC with unidirectional cooperation*—one transmitter knows the other's message and thus can enhance the achievable rate region.

The definition of degradedness of message sets can be further generalized to a K-user IFC: there are K messages W_t, $t = 1, \dots, K$, and W_t is to be decoded by receiver t while Transmitter t has a set of messages $\mathcal{M}_t$. The message sets are degraded if there exists a permutation $\{\sigma_k, k = 1, \dots, K\}$ of $\{1, 2, \dots, K\}$, such that

$$\mathcal{M}_{\sigma_1} \subset \cdots \subset \mathcal{M}_{\sigma_K}.$$

In general, the capacity region of an IFC is an open problem and is only known for certain classes of IFCs, which include the so-called *strong interference* channels that satisfy

$$\begin{aligned} I(X_1; Y_1 \mid X_2) &\le I(X_1; Y_2 \mid X_2) \\ I(X_2; Y_2 \mid X_1) &\le I(X_2; Y_1 \mid X_1), \end{aligned} \tag{6.2}$$

for all product distributions on the inputs X_1 and X_2. The capacity region in this case coincides with the capacity region of a compound IFC which is the union of two compound multiple access channels (MACs), as discovered by Ahlswede [1]. Maric, Yates and Kramer find the capacity region of strong IFCs with common information [68], and with degraded message sets [69]. An achievable region for IFCs with degraded message sets in a more general setting can be found in [30].

6.2.2 Gaussian IFCs

One of our main interests is the Gaussian IFC, in which the alphabets of inputs and outputs are real numbers and the outputs are linear combinations of input signals and white Gaussian noise. A Gaussian IFC is defined by

$$\begin{aligned} Y_1 &= X_1 + aX_2 + Z_1 \\ Y_2 &= bX_1 + X_2 + Z_2 \,, \end{aligned} \tag{6.3}$$

where a, b are real numbers, and Z_1, Z_2 are independent, zero-mean, unit-variance Gaussian random variables. Furthermore, the transmitters are subject to average power constraints:

$$\overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[X_{tn}^2] \le P_t \,, \quad t = 1, 2 \,.$$

The capacity region of the standard Gaussian IFC has been characterized in the following cases: (i) when $a = b = 0$ (trivial); (ii) either $a = 0$, $b \ge 1$ or $a \ge 1$, $b = 0$; and (iii) if $a^2 \ge 1$ and $b^2 \ge 1$, in which case the strong interference conditions in (6.2) are satisfied. The capacity of an IFC with strong interference is the set of (R_1, R_2) satisfying (see [43, 89])

$$0 \le R_1 \le \frac{1}{2}\log(1 + P_1) \tag{6.4a}$$

$$0 \le R_2 \le \frac{1}{2}\log(1 + P_2) \tag{6.4b}$$

$$0 \le R_1 + R_2 \le \frac{1}{2}\log(P_1 + a^2P_2 + 1) \tag{6.4c}$$

$$0 \le R_1 + R_2 \le \frac{1}{2}\log(b^2P_1 + P_2 + 1) \,. \tag{6.4d}$$

Note on terminology: when *either* $0 \le a^2 \le 1$ or $0 \le b^2 \le 1$ is satisfied, we say that the Gaussian IFC satisfies the weak interference condition. Achievable rate regions [24, 43, 87] and outer bounds [22, 24, 56, 88] are known for this scenario, but a characterization of the region is yet to be obtained. A recent outer bound by Kramer in [56] is given by (R_1, R_2) satisfying (6.4a)–(6.4b), and

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2}\log\left[\frac{(P_2+1)(P_1+a^2P_2+1)}{\min(a^2,1)P_2+1}\right] \\ R_1 + R_2 &\leq \frac{1}{2}\log\left[\frac{(P_1+1)(P_2+b^2P_1+1)}{\min(b^2,1)P_1+1}\right]. \end{aligned} \tag{6.5}$$

Let the capacity region of a Gaussian IFC-DMS (i.e., Figure 6.3) be denoted by $\mathcal{C}_{\text{GIFC}}^{T_1}$. In this chapter, we characterize the capacity region for the class of Gaussian weak interference channels, $\mathcal{C}_{\text{GIFC}}^{T_1}$, with $|b| \leq 1$ for any real valued a.

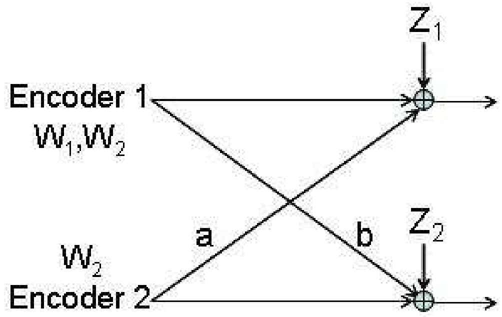


Figure 6.3: The Gaussian interference channel with degraded message sets.

6.2.3 Some intuitions on Gaussian IFC-DMS

Here we provide an intuitive view on the outer bound of Gaussian IFC-DMS, and set up the connection with its counterpart of Gaussian broadcast channel.

A straightforward outer bound $\mathcal{C}_{\text{GIFC}}^{T_1}$ for a Gaussian IFC-DMS is the capacity region of the Gaussian broadcast channel resulting from allowing full transmitter-side cooperation. One can derive an even tighter outer bound by using the following arguments:

i. Removing the interference link from Encoder 2 to Receiver 1 in Figure 6.3 does not enhance the overall capacity region, because Encoder 2 does not have any knowledge of message W_1, and thus no cooperation can be induced to improve the transmission rate R_1;

ii. Allowing full cooperation between the two encoders provides us with a new broadcast channel as shown in Figure 6.4, which has two transmit antennas and one receive antenna at each receiver and an individual power constraint at each antenna.

Using the existing literature on the capacity region of a Gaussian multi-antenna (MIMO) BC channel [104], dirty-paper coding (DPC) [25] optimizes this outer bound [101] [102]. If, in the dirty-paper coding strategy, W_1 is encoded first and W_2 second, the rates achieved are given by:

$$\mathcal{R}_{\text{DPC}}^{12} = \Bigg\{ (R_1, R_2) : R_1 \leq \frac{1}{2}\log\big(1+\alpha P_1\big)\,, \\ R_2 \leq \frac{1}{2}\log\left(\frac{1+P_1 b^2 + 2|b|\sqrt{(1-\alpha)P_1P_2}+P_2}{1+b^2\alpha P_1}\right), \text{for } 0 \leq \alpha \leq 1 \Bigg\}. \quad (6.6)$$

The DPC achievable region for the encoding sequence W_2, W_1 is

$$\mathcal{R}_{\text{DPC}}^{21} = \Bigg\{ (R_1, R_2) : R_1 \leq \frac{1}{2}\log\big(\frac{1+P_1}{1+(1-\alpha)P_1}\big)\,, \\ R_2 \leq \frac{1}{2}\log\big(1+P_1 b^2 + 2|b|\sqrt{(1-\alpha)P_1P_2}+P_2\big), \text{for } 0 \leq \alpha \leq 1 \Bigg\}. \quad (6.7)$$

It is not hard to verify that the capacity region of the Gaussian IFC-DMS in Figure 6.3 satisfies

$$\mathcal{C}_{\text{GIFC}}^{T_1} \subset \mathcal{R}_{\text{DPC}}^{12} \cup \mathcal{R}_{\text{DPC}}^{21} = \mathcal{R}_{\text{DPC}}^{21}\,.$$

We show later that the capacity region $\mathcal{C}_{\text{GIFC}}^{T_1}$ under weak interference is indeed equal to $\mathcal{R}_{\text{DPC}}^{12}$, and thus, even though the outer bound above is in general loose, it captures the intuition behind the optimal coding strategy for this channel.

6.3 Main results

In this section, we first obtain inner and outer bounds for the discrete-memoryless IFC-DMS, and then determine the capacity region of the Gaussian IFC-DMS with weak interference.

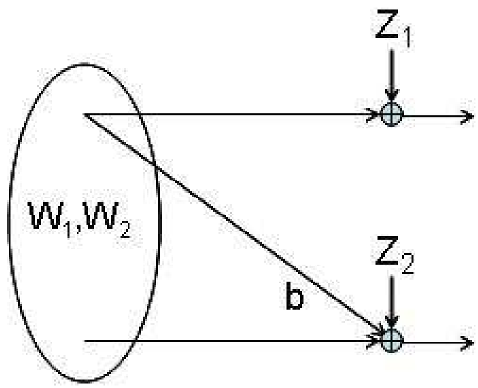


Figure 6.4: The Gaussian MIMO BC

6.3.1 Achievable regions and outer bounds for discrete-memoryless IFC-DMS

In this subsection, we provide achievable regions and outer bounds for general discrete memoryless IFC-DMS and then specialize to a class of IFC-DMS, whose capacity region is established.

We first present an achievable region for general discrete memoryless IFC-DMS.

Definition 6.3.1. *Define $\mathbb{R}_{in}$ to be the convex hull of all rate pairs (R_1, R_2) satisfying*

$$
\begin{aligned}
R_1 &\leq I(V;Y_1) - I(V;U,X_2)\\
R_2 &\leq I(U,X_2;Y_2)\,,
\end{aligned}
$$

over all probability distributions $p(x_1,x_2,u,v,y_1,y_2)$ that factor as

$$p(u,x_2)p(v \mid u,x_2)p(x_1 \mid v,u,x_2)p(y_1,y_2 \mid x_1,x_2)\,.$$

The following proposition gives the achievable region of IFC with Transmitter 1 knowing both messages as in Figure 6.2 using the Gel'fand-Pinsker coding scheme [40].

Proposition 6.3.1. *The capacity region of the discrete memoryless IFC-DMS in* (6.1) *satisfies*

$$\mathbb{R}_{in} \subset \mathcal{C}_{IFC}^{T_1}.$$

The proof of Proposition 6.3.1 follows that of the Gel'fand-Pinsker coding scheme in [40], thus it is omitted. Note that (U, X_2) are considered as the random parameters for the channel between X_1 and Y_1.

An outer bound is stated next.

Definition 6.3.2. *Define* $\mathbb{R}_o$ *to be the union of all rate pairs* (R_1, R_2) *satisfying*

$$\begin{aligned} R_1 &\le I(X_1; Y_1 \mid X_2) \\ R_2 &\le I(U, X_2; Y_2) \\ R_1 + R_2 &\le I(X_1; Y_1 \mid U, X_2) + I(U, X_2; Y_2)\,, \end{aligned} \tag{6.8}$$

over all probability distributions $p(x_1, x_2, u, y_1, y_2)$ *that factor as*

$$p(u, x_2) p(x_1 \mid u, x_2) p(y_1, y_2 \mid x_1, x_2)\,. \tag{6.9}$$

Theorem 6.3.2. *The capacity region of the discrete memoryless IFC-DMS in* (6.1) *satisfies*

$$\mathcal{C}_{IFC}^{T_1} \subset \mathbb{R}_o\,.$$

Proof. Theorem 6.3.2 can be proved by adapting the Körner-Marton's BC outer bound [55]. For a $(R_1, R_2, n, P_{e,1}^{(n)}, P_{e,2}^{(n)})$ code with decoding error $P_{e,i}^{(n)} \to 0$, as $n \to \infty$, we define the auxiliary random variable U by

$$U_i = (W_2, Y_1^{i-1}, Y_{2,i+1}^n, X_2^{i-1}, X_{2,i+1}^n)\,. \tag{6.10}$$

Applying Fano's inequality [26], for each message W_t, $t = 1, 2$, we have

$$H(W_t \mid Y_t^n) \le n R_t P_{e,t}^{(n)} + H(P_{e,t}^{(n)}) = n\varepsilon_t^{(n)}\,, \tag{6.11}$$

where $\varepsilon_t^{(n)} \to 0$, as $P_{e,t}^{(n)} \to 0$. Moreover, because Transmitter 2 has no information about the message W_1, X_2^n is independent of W_1, and the following relation holds

$$H(W_1 \mid W_2, X_2^n) = H(W_1)\,. \tag{6.12}$$

To prove the converse, we need the following lemma:

Lemma 6.3.3 ([27])**.** *For any random variable T, the following equality holds,*

$$\sum_{i=1}^{n} I(Y_{2,i+1}^n; Y_{1,i} \mid Y_1^{i-1}, T) = \sum_{i=1}^{n} I(Y_1^{i-1}; Y_{2,i} \mid Y_{2,i+1}^n, T)\,.$$

Define the time-sharing random variable T, where $T = i \in \{1, \ldots, n\}$ with probability $1/n$, and

$$\begin{aligned}
X_1 &\triangleq X_{1,T} \\
X_2 &\triangleq X_{2,T} \\
Y_1 &\triangleq Y_{1,T} \\
Y_2 &\triangleq Y_{2,T} \\
\tilde{U} &\triangleq (U_T, T)\,.
\end{aligned}$$

First we prove the outer bounds for R_1 and R_2 in (6.8). We have

$$\begin{aligned}
nR_1 &\leq I(W_1; Y_1^n \mid X_2^n) + n\varepsilon_1^{(n)} \qquad (6.13)\\
&= \sum_{i=1}^{n} I(W_1; Y_{1,i} \mid Y_1^{i-1}, X_2^n) + n\varepsilon_1^{(n)} \\
&\leq \sum_{i=1}^{n} \big[H(Y_{1,i} \mid X_{2,i}) - H(Y_{1,i} \mid W_1, X_2^n, Y_1^{i-1}, X_{1,i})\big] + n\varepsilon_1^{(n)} \\
&= \sum_{i=1}^{n} I(X_{1,i}; Y_{1,i} \mid X_{2,i}) + n\varepsilon_1^{(n)} \\
&= nI(X_{1,T}; Y_{1,T} \mid X_{2,T}, T) + n\varepsilon_1^{(n)} \\
&\leq nI(X_1; Y_1 \mid X_2) + n\varepsilon_1^{(n)}\,,
\end{aligned}$$

and

$$
\begin{aligned}
nR_2 &\le I(W_2;Y_2^n) + n\varepsilon_2^{(n)} && (6.14)\\
&= \sum_{i=1}^{n} I(W_2;Y_{2,i} \mid Y_{2,i+1}^n) + n\varepsilon_2^{(n)}\\
&\le \sum_{i=1}^{n} I(W_2,Y_{2,i+1}^n;Y_{2,i}) + n\varepsilon_2^{(n)}\\
&\le \sum_{i=1}^{n} I(U_i,X_{2,i};Y_{2,i}) + n\varepsilon_2^{(n)}\\
&= nI(U_T,X_{2,T};Y_{2,T} \mid T) + n\varepsilon_2^{(n)}\\
&\le nI(U_T,T,X_{2,T};Y_{2,T}) + n\varepsilon_2^{(n)}\\
&= nI(\tilde{U},X_2;Y_2) + n\varepsilon_2^{(n)}\,,
\end{aligned}
$$

where (6.13) and (6.14) are from Fano's inequality in (6.11).

Next, we prove the outer bound for the sum rate $R_1 + R_2$ in (6.8). We have

$$
\begin{aligned}
n(R_1+R_2) &\le I(W_1;Y_1^n \mid W_2,X_2^n) + I(W_2;Y_2^n) + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)} && (6.15a)\\
&\le \sum_{i=1}^{n} \big[I(W_1;Y_{1,i} \mid W_2,X_2^n,Y_1^{i-1})\\
&\qquad + I(W_2;Y_{2,i} \mid Y_{2,i+1}^n)\big] + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)} && (6.15b)\\
&= \sum_{i=1}^{n} \big[I(W_1,Y_{2,i+1}^n;Y_{1,i} \mid W_2,X_2^n,Y_1^{i-1})\\
&\qquad - I(Y_{2,i+1}^n;Y_{1,i} \mid W_1,W_2,X_2^n,Y_1^{i-1})\\
&\qquad + I(W_2,Y_{2,i+1}^n,X_2^n;Y_{2,i}) - I(X_2^n;Y_{2,i} \mid W_2,Y_{2,i+1}^n)\\
&\qquad - I(Y_{2,i+1}^n;Y_{2,i})\big] + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)}\,. && (6.15c)
\end{aligned}
$$

Note (6.15a) is due to Fano's inequality in (6.11) and the conditional entropy relation in (6.12); the first two terms in (6.15c) are from the first term in (6.15b) and the third, fourth and fifth terms are from the second term in (6.15b). Since mutual information is nonnegative, by dropping the second, fourth and fifth terms in (6.15c), we obtain

$$
\begin{aligned}
n(R_1+R_2) &\le \sum_{i=1}^{n}\big[I(W_1, Y_{2,i+1}^n; Y_{1,i} \mid W_2, X_2^n, Y_1^{i-1}) \\
&\qquad + I(W_2, Y_{2,i+1}^n, X_2^n; Y_{2,i})\big] + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)} \\
&\le \sum_{i=1}^{n}\big[I(W_1; Y_{1,i} \mid W_2, Y_1^{i-1}, Y_{2,i+1}^n, X_2^n) \\
&\qquad + I(Y_{2,i+1}^n; Y_{1,i} \mid W_2, Y_1^{i-1}, X_2^n) \\
&\qquad + I(W_2, Y_1^{i-1}, Y_{2,i+1}^n, X_2^n; Y_{2,i}) \\
&\qquad - I(Y_1^{i-1}; Y_{2,i} \mid W_2, Y_{2,i+1}^n, X_2^n)\big] + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)} && (6.15\text{d}) \\
&\le \sum_{i=1}^{n}\big[I(W_1; Y_{1,i} \mid U_i, X_{2,i}) + I(U_i, X_{2,i}; Y_{2,i})\big] + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)} && (6.15\text{e}) \\
&\le \sum_{i=1}^{n}\big[I(X_{1,i}; Y_{1,i} \mid U_i, X_{2,i}) + I(U_i, X_{2,i}; Y_{2,i})\big] + n\varepsilon_1^{(n)} + n\varepsilon_2^{(n)} && (6.15\text{f}) \\
&= nI(X_{1,T}; Y_{1,T} \mid U_T, X_{2,T}, T) + n\varepsilon_1^{(n)} \\
&\qquad + nI(U_T, X_{2,T}; Y_{2,T} \mid T) + n\varepsilon_2^{(n)} \\
&\le nI(X_1; Y_1 \mid \tilde{U}, X_2) + n\varepsilon_1^{(n)} + nI(\tilde{U}, X_2; Y_2) + n\varepsilon_2^{(n)}
\end{aligned}
$$

In this calculation, the second and fourth terms in (6.15d) are equal from Lemma 6.3.3; (6.15e) is obtained by using the auxiliary random variable U_i defined in (6.10); and (6.15f) is true because $(W_1, U_i) \Rightarrow (X_{1,i}, X_{2,i}) \Rightarrow (Y_{1,i}, Y_{2,i})$, for all $1 \le i \le n$. As the auxiliary random variable $\tilde{U}$ includes the time-sharing random variable T and is an element in the set of auxiliary random variables defined for $\mathbb{R}_o$, it is straightforward to show $\mathbb{R}_o$ is convex. □

Both Proposition 6.3.1 and Theorem 6.3.2 hold for the general IFCs. However, as seen, the achievable region obtained in Proposition 6.3.1 does not, in general, meet the outer bound in Theorem 6.3.2.

Next we investigate the scenarios under which the capacity region of IFC-DMS can be obtained under additional assumptions.

Definition 6.3.3. *Define the rate region $\mathbb{R}_*$ to be the union of all rate pairs (R_1, R_2)*

satisfying

$$
\begin{aligned}
R_1 &\leq I(X_1; Y_1 \mid U, X_2) \\
R_2 &\leq I(U, X_2; Y_2)\,,
\end{aligned}
\tag{6.16}
$$

over all probability distributions $p(x_1, x_2, u, y_1, y_2)$ *that factor as*

$$
p(u, x_2)p(x_1 \mid u, x_2)p(y_1, y_2 \mid x_1, x_2)\,.
$$

It is not difficult to see that $\mathbb{R}_*$ is a subset of $\mathbb{R}_o$, namely, $\mathbb{R}_* \subset \mathbb{R}_o$. As shown in Figure 6.5, for a fixed auxiliary random variable U, since $I(X_1; Y_1 \mid U, X_2) \leq I(X_1; Y_1 \mid X_2)$, the rate region defined by (6.8) in $\mathbb{R}_o$ corresponds to the area $OABCD$ in Figure 6.5 while the rate region defined by (6.16) in $\mathbb{R}_*$ corresponds to the shaded rectangle $OABE$ in Figure 6.5.

Under the following assumption on the channel, we can show that $\mathbb{R}_o = \mathbb{R}_*$.

Assumption 6.3.1.

$$
I(U; Y_2 \mid X_2) \leq I(U; Y_1 \mid X_2) \tag{6.17}
$$

is satisfied for all auxiliary random variables U*, such that the probability distribution* $p(x_1, x_2, u, y_1, y_2)$ *can factor as* $p(u, x_2)p(x_1 \mid u, x_2)p(y_1, y_2 \mid x_1, x_2)$.

Proposition 6.3.4. *Under Assumption 6.3.1,*

$$
\mathbb{R}_o = \mathbb{R}_*\,,
$$

thus the capacity region of the discrete memoryless IFC-DMS satisfying Assumption 6.3.1 satisfies

$$
\mathcal{C}_{IFC}^{T_1} \subset \mathbb{R}_*\,.
$$

Proof. Because the auxiliary random variable U is over a similar set of probability distributions, it is enough to compare rate regions defined by inequalities in (6.8) and (6.16) given a probability distribution $p(x_1, x_2, u, y_1, y_2)$ factored as (6.9). As shown in Figure 6.5, for a given $p(x_1, x_2, u, y_1, y_2)$, the rate region defined by (6.8)

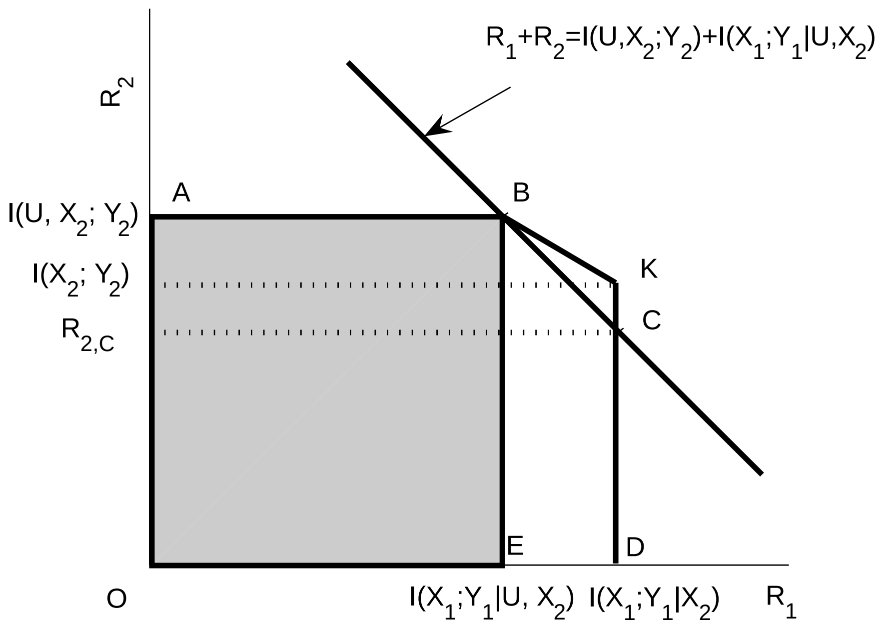


Figure 6.5: Comparison between rate regions $\mathbb{R}_o$ and $\mathbb{R}_*$ for a fixed probability distribution $p(x_1, x_2, u, y_1, y_2)$ factored as (6.9).

corresponds to the area $OABCD$ and the rate region defined by (6.16) corresponds to the area $OABE$. Thus clearly $\mathbb{R}_* \subset \mathbb{R}_o$.

To show $\mathbb{R}_o \subset \mathbb{R}_*$, it is enough to show that, for the given $p(x_1, x_2, u, y_1, y_2)$, the point C is in $\mathbb{R}_*$ due to the convexity of $\mathbb{R}_*$. Under Assumption 6.3.1, the R_2-coordinate of the point C, $R_{2,C}$, satisfies

$$\begin{aligned} R_{2,C} &= I(U, X_2; Y_2) + I(X_1; Y_1 \mid U, X_2) - I(X_1; Y_1 \mid X_2) \\ &= I(U, X_2; Y_2) + H(Y_1 \mid U, X_2) - H(Y_1 \mid X_2) \\ &= I(X_2; Y_2) + I(U; Y_2 \mid X_2) - I(U; Y_1 \mid X_2) \\ &\leq I(X_2; Y_2)\,. \end{aligned}$$

Let U be a constant, and denote the corresponding point in (6.16) as K, then we have

$$(R_{1,K}, R_{2,K}) = \big(I(X_1; Y_1 \mid X_2), I(X_2; Y_2)\big) \in \mathbb{R}_*\,.$$

Since $R_{2,C} \leq R_{2,K}$, C is in the area $OABKD$, thus $C \in \mathbb{R}_*$. Therefore $\mathbb{R}_o \subset \mathbb{R}_*$. $\square$

Next we show $\mathbb{R}_*$ is also achievable under a further assumption on the channel together with Assumption 6.3.1, thus the capacity region is established.

Assumption 6.3.2. *For an IFC,*

$$I(X_2; Y_2) \leq I(X_2; Y_1) \tag{6.18}$$

is satisfied over all input distributions to the channel $p(x_1, x_2)$.

Theorem 6.3.5. *The capacity region of discrete memoryless IFC-DMS satisfying both Assumption 6.3.1 and Assumption 6.3.2 is*

$$\mathcal{C}_{IFC}^{T_1} = \mathbb{R}_*\,.$$

Proof. First note that combining (6.17) and (6.18), it is easy to see the IFC-DMS must satisfy

$$I(U, X_2; Y_2) \leq I(U, X_2; Y_1)\,, \tag{6.19}$$

for all probability distributions $p(x_1, x_2, u, y_1, y_2)$ that factor as

$$p(u, x_2)p(x_1 \mid u, x_2)p(y_1, y_2 \mid x_1, x_2)\,.$$

Under (6.19), we have the following coding scheme based on superposition coding [26]:

Code Generation: Fix $p(u, x_2)$, and generate 2^{nR_2} independent codewords of length n at random according to the distribution $\prod_{i=1}^{n} p(u_i, x_{2,i})$, for message $w_2 \in \{1, \ldots, 2^{nR_2}\}$. For each codeword $(U^n(w_2), X_2^n(w_2))$, generate 2^{nR_1} independent codewords $X_1^n(w_1, w_2)$ according to $\prod_{i=1}^{n} p(x_{1,i} \mid u_i(w_2), x_{2,i}(w_2))$, with $w_1 \in \{1, \ldots, 2^{nR_1}\}$.

Encoding: Encoder 2 transmits $X_2^n(W_2)$. Since Encoder 1 knows both messages, it sends $X_1^n(W_1, W_2)$.

Decoding: Receiver 2 determines the unique $\hat{\hat{W}}_2$ such that $(U^n(\hat{\hat{W}}_2), X_2^n(\hat{\hat{W}}_2), Y_2^n)$ is jointly typical. If there are none such or more than one such, an error is declared. Receiver 1 determines the unique $(\hat{W}_1, \hat{W}_2)$ such that

$$(X_1^n(\hat{W}_1, \hat{W}_2), X_2^n(\hat{W}_2), U^n(\hat{W}_2), Y_1^n)$$

is jointly typical. Again, if there are none such or more than one such, an error is declared.

It is easy to see that the probability of error at Receiver 2 tends to zero, as $n \to \infty$, if $R_2 \le I(U, X_2; Y_2)$, and Receiver 1 can decode W_2 successfully, as $n \to \infty$, provided $R_2 \le I(U, X_2; Y_1)$. Under (6.19),

$$R_2 \le I(U, X_2; Y_2) \le I(U, X_2; Y_1)\,.$$

Thus, Receiver 1 can decode W_2 as long as Receiver 2 can do so. With the error probability of W_2 tending to zero, as $n \to \infty$, the error probability of W_1 at Receiver 1 goes to zero, provided $R_1 \le I(X_1; Y_1 \mid U, X_2)$. The above analysis shows that both receivers can decode with the total probability of error tending to zero, if (6.16) is satisfied. Hence there exists a sequence of good codes with error probability tending to 0. $\square$

6.3.2 Gaussian IFC-DMS with weak interference

Next we investigate Gaussian IFC-DMS with weak interference when $|b| \leq 1$.

Note for a general IFC that the two receivers cannot cooperate, the capacity region is the same as the one with the same marginal output $p(y_1 \mid x_1, x_2)$, $p(y_2 \mid x_1, x_2)$. The same result holds for Gaussian IFC-DMS as stated by the following lemma.

Lemma 6.3.6. *The capacity region of a Gaussian IFC-DMS given by* (6.3) *when* $|b| \leq 1$*, is the same as that of a Gaussian IFC-DMS defined as*

$$\begin{aligned} \tilde{Y}_1 &= Y_1' + aX_2 \\ \tilde{Y}_2 &= bY_1' + X_2 + \tilde{Z}\,, \end{aligned} \tag{6.20}$$

where

$$Y_1' = X_1 + Z_1\,,$$

and $\tilde{Z}$ *is Gaussian distributed with variance* $1-b^2$ *and independent of* Z_1*, i.e., given* X_2*,*

$$X_1 \Rightarrow \tilde{Y}_1 \Rightarrow \tilde{Y}_2\,. \tag{6.21}$$

Proof. This follows directly from the arguments in [26] (10, Pg. 454) or [88]. □

Before determining its capacity region of Gaussian IFC-DMS, we first present an outer bound tailored for Gaussian IFC-DMS with weak interference.

Lemma 6.3.7. *The capacity region of weak interference Gaussian IFC-DMS with* $|b| \leq 1$ *satisfies*

$$\mathcal{C}_{GIFC}^{T_1} \subset \mathbb{R}_*\,.$$

Proof. According to the result in Lemma 6.3.6, when $|b| \leq 1$, the capacity region of any Gaussian IFC-DMS is equal to that of a Gaussian IFC-DMS satisfying that, given X_2

$$X_1 \Rightarrow Y_1 \Rightarrow Y_2\,. \tag{6.22}$$

Thus it is enough to prove the outer bound for the Gaussian IFC-DMS satisfying (6.22).

Afterward, the key is to identify the auxiliary random variable. Define the time-sharing random variable T, where $T = i \in \{1, \ldots, n\}$ with probability $1/n$, and $X_1 \triangleq X_{1,T}$, $X_2 \triangleq X_{2,T}$, $Y_1 \triangleq Y_{1,T}$, $Y_2 \triangleq Y_{2,T}$, $\tilde{U} \triangleq (U_T, T)$.

For any $(R_1, R_2, n, P_{e,1}^{(n)}, P_{e,2}^{(n)})$ code with decoding error $P_{e,i}^{(n)} \to 0$, as $n \to \infty$, we have

$$
\begin{aligned}
nR_1 &= H(W_1) \\
&\leq I(W_1; Y_1^n) + n\varepsilon_1^{(n)} && \text{(6.23a)} \\
&\leq I(W_1; Y_1^n \mid W_2, X_2^n) + n\varepsilon_1^{(n)} && \text{(6.23b)} \\
&= \sum_{i=1}^{n} I(W_1; Y_{1,i} \mid W_2, Y_1^{i-1}, X_2^n) + n\varepsilon_1^{(n)}\,, && \text{(6.23c)}
\end{aligned}
$$

where (6.23a) is due to Fano's inequality (6.11); (6.23b) follows from (6.12); (6.23c) is due to the chain rule for mutual information.

Defining $U_i = (W_2, Y_1^{i-1}, X_2^{i-1})$, we obtain

$$\begin{aligned}
nR_1 &\leq \sum_{i=1}^{n} I(W_1; Y_{1,i} \mid U_i, X_{2,i}, X_{2,i+1}^n) + n\varepsilon_1^{(n)} \\
&= \sum_{i=1}^{n} \big[h(Y_{1,i} \mid U_i, X_{2,i}, X_{2,i+1}^n) \\
&\qquad - h(Y_{1,i} \mid U_i, X_{2,i}, X_{2,i+1}^n, W_1)\big] + n\varepsilon_1^{(n)} \\
&\leq \sum_{i=1}^{n} \big[h(Y_{1,i} \mid U_i, X_{2,i}) \\
&\qquad - h(Y_{1,i} \mid U_i, X_{2,i}, X_{2,i+1}^n, W_1, X_{1,i})\big] + n\varepsilon_1^{(n)} \qquad (6.23d) \\
&= \sum_{i=1}^{n} \big[h(Y_{1,i} \mid U_i, X_{2,i}) \\
&\qquad - h(Y_{1,i} \mid U_i, X_{1,i}, X_{2,i})\big] + n\varepsilon_1^{(n)} \qquad (6.23e) \\
&= \sum_{i=1}^{n} \big[I(X_{1,i}; Y_{1,i} \mid U_i, X_{2,i})\big] + n\varepsilon_1^{(n)} \\
&= nI(X_{1,T}; Y_{1,T} \mid U_T, X_{2,T}, T) + n\varepsilon_1^{(n)} \\
&= nI(X_1; Y_1 \mid \tilde{U}, X_2) + n\varepsilon_1^{(n)}
\end{aligned}$$

where (6.23d) follows from the fact that entropy decreases by adding conditionals, and (6.23e) holds since $W_1 \Rightarrow (X_{1,i}, X_{2,i}) \Rightarrow Y_{1,i}$.

The bound of R_2 in (6.16) can be derived as follows:

$$\begin{aligned}
nR_2 &= H(W_2) \\
&\leq I(W_2; Y_2^n) + n\varepsilon_2^{(n)} \qquad (6.24a) \\
&= \sum_{i=1}^{n} I(W_2; Y_{2,i} \mid Y_2^{i-1}) + n\varepsilon_2^{(n)} \\
&\leq \sum_{i=1}^{n} \big[h(Y_{2,i}) - h(Y_{2,i} \mid Y_2^{i-1}, W_2, X_2^i)\big] + n\varepsilon_2^{(n)} \qquad (6.24b) \\
&= nI(Y_{2,T}; U_T, X_{2,T} \mid T) + n\varepsilon_2^{(n)} \\
&\leq nI(Y_2; \tilde{U}, X_2) + n\varepsilon_2^{(n)},
\end{aligned}$$

where (6.24a) follows from Fano's inequality in (6.11). Due to (6.22), given X_2,

$$Y_{2,i} \Rightarrow Y_1^{i-1} \Rightarrow Y_2^{i-1}$$

because conditioning on (X_2^{i-1}, Y_1^{i-1}), Y_2^{i-1} is independent of other random variables (e.g., $Y_{2,i}$), or,

$$I(Y_{2,i}; Y_1^{i-1} \mid W_2, X_2^i) \geq I(Y_{2,i}; Y_2^{i-1} \mid W_2, X_2^i)\,.$$

Thus the conditional entropy in (6.24b) satisfies

$$\begin{aligned} h(Y_{2,i} \mid Y_2^{i-1}, W_2, X_2^i) &\geq h(Y_{2,i} \mid Y_1^{i-1}, W_2, X_2^i) \\ &= h(Y_{2,i} \mid U_i, X_{2,i}) \end{aligned} \tag{6.25}$$

Combining (6.24b) and (6.25), we obtain the result in Lemma 6.3.7. Moreover, since the time-sharing random variable is included in $\tilde{U}$, it is straightforward to show the outer bound is a convex set applying the same arguments in the proof for Theorem 6.3.2. □

Theorem 6.3.8. *The capacity region $\mathcal{C}_{GIFC}^{T_1}$ of the Gaussian IFC-DMS with Transmitter 1 knowing both messages, when $|b| \leq 1$, is the set of all rate pairs (R_1, R_2) such that, for $0 \leq \alpha \leq 1$,*

$$R_1 \leq \frac{1}{2}\log\left(1 + \alpha P_1\right) \tag{6.26}$$

$$\begin{aligned} R_2 &\leq \frac{1}{2}\log\left(1 + \frac{h\Sigma h^t}{1 + b^2\alpha P_1}\right) \\ &= \frac{1}{2}\log\left(\frac{1 + P_1 b^2 + 2|b|\sqrt{(1-\alpha)P_1P_2} + P_2}{1 + b^2\alpha P_1}\right) \end{aligned} \tag{6.27}$$

In (6.27), h is the vector $[b\ 1]$, and Σ is a 2×2 covariance with diagonal elements equaling $(1-\alpha)P_1$ and P_2 respectively.

Proof. Achievability: The proof of achievability of this rate utilizes the dirty-paper coding strategy. First, we generate a codebook of 2^{nR_2} codewords according to

$\mathcal{N}(0, \Sigma)$, where Σ is the covariance between Transmitters 1 and 2. Transmitter 1 devotes a fraction $(1-\alpha)$ of its power P_1 to the transmission of W_2, while Transmitter 2 devotes its entire power P_2 to this effort. This yields a covariance of the form

$$\Sigma = \begin{bmatrix} (1-\alpha)P_1 & \gamma \\ \gamma & P_2 \end{bmatrix} \tag{6.28}$$

The effective interference seen by Receiver 1 is a combination of the signals communicated from both Transmitters 1 and 2. Since Transmitter 1 knows the exact realization of the message $w_2 \in W_2$, it has non-causal side information on the interference and can completely cancel it out, achieving a rate

$$R_1 = \frac{1}{2}\log\left(1 + \alpha P_1\right) , \tag{6.29}$$

by using a Gaussian codebook with codewords that are correlated with the interference. At Receiver 2, this Gaussian codebook for W_1 is perceived as additive interference, hence achieving a rate

$$R_2 = \frac{1}{2}\log\left(1 + \frac{h\Sigma h^t}{1 + b^2\alpha P_1}\right) . \tag{6.30}$$

Maximizing R_2 over $|\gamma|^2 \leq (1-\alpha)P_1P_2$ (i.e., keeping Σ positive definite), we find that R_2 attains its maximum when $\gamma = \sqrt{(1-\alpha)P_1P_2}$. and (6.27) can be achieved.

Converse: Since for Gaussian IFC-DMS with $|b| \leq 1$, both Lemma 6.3.6 and Lemma 6.3.7 hold. Thus it remains to prove the optimality of Gaussian input for the Gaussian IFC-DMS redefined as

$$\begin{aligned} Y_1 &= X_1 + aX_2 + Z_1 \\ Y_2 &= b(X_1 + Z_1) + X_2 + Z' , \end{aligned} \tag{6.31}$$

where Z' is a Gaussian distributed random variable with variance $1 - b^2$ and independent of Z_1. For this, we need following lemmas:

Lemma 6.3.9 (Lemma 1 in [99])**.** *Let $X_1, X_2, \ldots, X_k$ be an arbitrary set of zero-mean random variables with covariance matrix K. Let S by any subset of* $\{1, 2, \ldots, k\}$

and $\bar{S}$ be its complement. Then

$$h(X_S \mid X_{\bar{S}}) \le h(X_S^* \mid X_{\bar{S}}^*)\,,$$

where

$$(X_1^*, X_2^*, \ldots, X_k^*) \sim N(0, K)\,.$$

First we note that

$$\begin{aligned} h(Y_1 \mid U, X_2) &\ge h(Y_1 \mid U, X_1, X_2) \\ &= h(Y_1 \mid X_1, X_2) \\ &= \frac{1}{2}\log(2\pi e)\,, \end{aligned}$$

and on the other hand,

$$\begin{aligned} h(Y_1 \mid U, X_2) &= h(X_1 + aX_2 + Z_1 \mid U, X_2) \\ &= h(X_1 + Z_1 \mid U, X_2) \\ &\le h(X_1 + Z_1) \\ &\le \frac{1}{2}\log\big(2\pi e(1 + P_1)\big)\,. \end{aligned}$$

Without loss of generality, we assume that

$$h(Y_1 \mid U, X_2) = \frac{1}{2}\log\big(2\pi e(1 + \alpha P_1)\big)\,, \tag{6.32}$$

for some $\alpha \in [0, 1]$. Thus, we obtain

$$\begin{aligned} I(X_1; Y_1 \mid U, X_2) &= h(Y_1 \mid U, X_2) - h(Y_1 \mid U, X_1, X_2) \\ &= \frac{1}{2}\log(1 + \alpha P_1)\,. \end{aligned} \tag{6.33}$$

On the other hand, by Lemma 6.3.9, the conditional entropy is upper bounded by the Gaussian variables with the same covariance matrix, thus,

$$\begin{aligned} h(Y_1 \mid U, X_2) &\le h(X_1 + Z_1 \mid X_2) \\ &= \frac{1}{2} \log\Big(2\pi e\big(1 + (1-\rho^2)P_1\big)\Big) . \end{aligned} \tag{6.34}$$

Note

$$\rho \triangleq \frac{\mathbb{E}[X_1 X_2]}{\sqrt{P_1 P_2}}$$

denoting the correlation coefficient, thus combining with (6.32)

$$\alpha \le 1 - \rho^2 . \tag{6.35}$$

Therefore,

$$\begin{aligned} h(Y_2) &= h(bX_1 + X_2 + Z_2) \\ &\le \frac{1}{2} \log\Big(2\pi e(1 + b^2 P_1 + P_2 + 2b\rho\sqrt{P_1 P_2}\Big) \\ &\le \frac{1}{2} \log\Big(2\pi e(1 + b^2 P_1 + P_2 \\ &\qquad\qquad + 2b\sqrt{(1-\alpha) P_1 P_2})\Big) . \end{aligned} \tag{6.36}$$

Next, we need to bound $h(Y_2 \mid X_2, U)$. According to (6.31), Y_2 is a degraded version of Y_1 conditioning on X_2. By the entropy power inequality (EPI) [26],

$$\begin{aligned} 2^{2h(Y_2|X_2,U)} &\ge 2^{2h(bY_1|X_2,U)} + 2^{2h(Z')} \\ &= b^2 2^{2h(Y_1|X_2,U)} + 2\pi e(1 - b^2) \\ &= 2\pi e(1 + b^2 \alpha P_1) , \end{aligned}$$

which yields

$$h(Y_2 \mid X_2, U) \ge \frac{1}{2} \log\Big(2\pi e(1 + b^2 \alpha P_1)\Big) . \tag{6.37}$$

Finally, we combine (6.36) and (6.37), to obtain

$$
\begin{aligned}
& I(X_2, U; Y_2) \\
& = h(Y_2) - h(Y_2 \mid X_2, U) \\
& \le \frac{1}{2} \log\left(\frac{1 + b^2 P_1 + P_2 + 2b\sqrt{(1-\alpha) P_1 P_2}}{1 + b^2 \alpha P_1} \right). \qquad (6.38)
\end{aligned}
$$

Since (6.33) and (6.38) are similar to (6.26) and (6.27), the optimality of Gaussian inputs is established for the Gaussian IFC-DMS redefined in (6.31) and its capacity region is obtained.

By Lemma 6.3.6, the capacity region of the original Gaussian IFC-DMS is equal to the region defined by (6.26) and (6.27) as well. The proof is complete. □

6.4 Numerical results

In this section, we use numerical results to compare the capacity region of Gaussian IFC-DMS with achievable rate regions of Gaussian BCs and the outer bound of Gaussian IFCs.

We consider a symmetric Gaussian IFC with $P_1 = P_2 = 6$ and $a^2 = b^2 = 0.3$. The rate units are bits per channel use. In Figure 6.6, we compare the capacity region of Gaussian IFC-DMS $\mathcal{C}_{\mathrm{GIFC}}^{T_1}$ with the dirty-paper coding regions $\mathcal{R}_{\mathrm{DPC}}^{12}$ and $\mathcal{R}_{\mathrm{DPC}}^{21}$, and the outer bound of Gaussian IFCs in (6.5). We observe that the capacity region of Gaussian IFC-DMS is strictly larger than the outer bound of Gaussian IFCs, and the gap between these two indicates the performance improvement by allowing encoders to cooperate partially. The point at $\alpha = 0$ corresponds to full cooperation between two encoders in Gaussian IFC-DMS to transmit the message W_2, and it meets the capacity region of Gaussian BCs. The corresponding rate for W_2 is equal to 1.971 bps, while Gaussian IFC without cooperation can only achieve 1.404 bps.

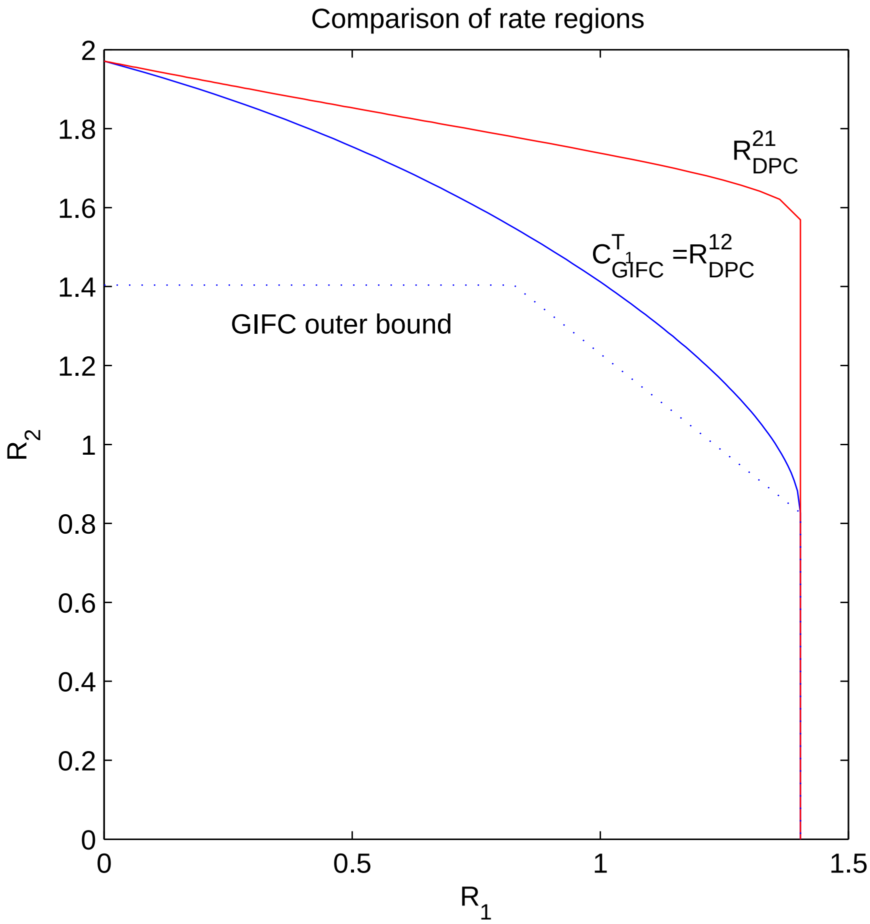


Figure 6.6: The capacity region of Gaussian IFC-DMS with $P_1 = P_2 = 6$, $a^2 = b^2 = 0.3$, two DPC rate regions of Gaussian BC in Figure 6.4, and the outer bound of Gaussian IFC in [56].

Chapter 7

Conclusion and Future Work

In this dissertation, we have studied several problems arising in the networked control systems, i.e., deploying sensor networks in control systems. In this chapter, we conclude this dissertation and point out possible further extensions for future research.

7.1 Sensor querying

In Chapter 3 and Chapter 4, we have studied a centralized control system enabled with sensor querying using a general framework of partially observable Markov decision processes (POMDP), and then specialized to cases including a hierarchical querying model (Chapter 3) and a stochastic linear system (Chater 4). The main results are

- A characterization of existence of ergodic control for finite state POMDPs with hierarchical information structure by the average cost optimality equation (ACOE);
- A necessary and sufficient condition for stability of a stochastic linear system;
- A full characterization of optimal control of LQG problem under sensor querying: the separation principle holds, namely, the optimal controller is a feedback control based on current state estimation while the optimal estimator is a switching Kalman filter where the switching is determined by solving an ACOE for sensor querying.

As sensor querying brings a new dimension of design into networked control systems, there are a number of interesting theoretical problems and practical implementation challenges that need to be further addressed.

7.1.1 Sensor querying and finite rate channel

In our work, data at each sensor is assumed to be transmitted perfectly once it is queried. In a more practical setup, in which a communication channel with finite data rate is associated with each sensor query, an interesting problem is that, under what conditions for these channels, the stability of the system can be preserved. Even in the simplest form, with the query variable is constant $Q_t = q$ for all t, this question is answered very recently by Wong and Brockett [105, 106], Tatikonda and Mitter [97, 98], Elia and Mitter [34], Nair and Evans [74], Liberzon [63]. A well-known result among them is that, if and only if (C, A) is detectable and the rate R satisfies

$$\sum_k \max\Big\{\log\big(\lambda_k(A)\big), 0\Big\} < R\,, \tag{7.1}$$

where $\lambda_k(A)$ is the k-th eigenvalue of the square matrix A. Stability of nonlinear control systems is further studied in [75] and [64].

With sensor querying, the question is, under what conditions for R_q, $q = 1, 2, \ldots, Q$, we can obtain the estimate of X_t, $\hat{X}_t$ via sensor querying, such that $\mathbb{E}[(X_t - \hat{X}_t)^2]$ is bounded. It is challenging not only to determine the optimal querying rule, but also to design coding schemes for sensors. Moreover, if the channel of each sensor is not a perfect digital link with length R_q, e.g., it is an erasure channel with certain erasure probability p_q, what is the coding scheme and corresponding sensor querying rule? Does the separation principle hold for the channel coding and sensor querying?

CEO problem with sensor querying

A feature of sensor querying can be added to a famous problem in multi-terminal information theory, the CEO problem [10, 77, 78, 103]. The CEO problem is a distributed source coding problem, in which there are N employees and one CEO. Each of employees observes a copy (Y_t^q) of the source (X_t), compresses and reports it to the CEO via an error-free channel with rate R_q. The CEO wants to reconstruct the

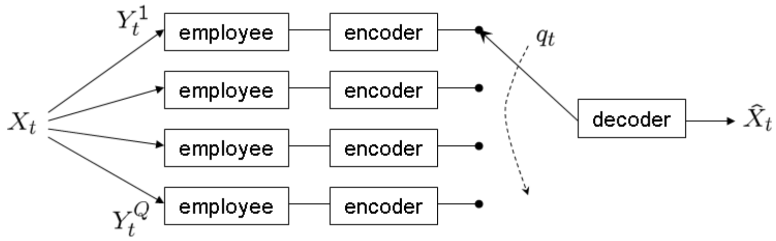


Figure 7.1: A variant of CEO problem with sensor querying

estimates of X_t such that the estimation variance is minimized. A distortion region can be characterized in term of the rate tuple and the Gaussian CEO problem has been fully solved by Oohama in [78].

Now since it is difficult to have a dedicated link between each employee and CEO, it is of practical importance to consider the CEO problem under sensor querying, namely, only one employee can report to CEO at certain rate at a time. A characterization of the distortion region or even an asymptotic result as $N \to \infty$ is particularly intriguing.

Sensor querying and system identification, learning

Estimation, control and identification problems, the latter arising when the system dynamics contain unknown parameters, can be addressed via the adaptive control of POMDPs Sensor querying adds a new dimension to recursive identification techniques [66], as well the study of the Robbins-Monro type algorithms for adaptive estimation problems in stochastic systems, via stochastic approximation methods [60,62]. For the standard stochastic algorithm, the estimate of certain unknown parameter is updated as

$$X_{t+1} = X_t + \epsilon_t(Y_t)\,,$$

where $Y_t = f(X_t, N_t)$ and N_t is some noise. Under some mild conditions, X_t can be proved to converge to x^* almost surely as $t \to \infty$ and $f(x^*) = 0$.

An interesting extension of sensor querying to such a setup is to consider that there are a number of system outputs Y_t^q available for the estimation with $Y_t^q = f_q(X_t, N_t)$ and we want to design a querying rule to achieve the convergence of the algorithm, namely X_t can converge to the intended value as well. Clearly, such a querying rule exists if one of the outputs satisfies the condition for convergence.

However, it is possible to obtain the convergence of the stochastic approximation algorithm even if each output is "weak" but "strong" cumulatively. Moreover, sensor querying might potentially be able to improve the convergence rate.

7.2 Optimal power allocation in a time-varying environment

In Chapter 5, we have studied optimal power allocation in a wireless fading channel concerning both queueing delay and power efficiency. Under a fast channel variation assumption, i.e., if the channel state changes much faster than the queueing dynamics, we have taken a heavy traffic analysis and associated a monotone cost function with the limiting queue-length process. In the heavy traffic limit, the discrete-time MDP problem is transformed into a continuous-time controlled diffusion problem. For the diffusion limit, we have first shown the existence of the optimal stationary Markov policy, and then shown that this is a channel-state based threshold policy. In other words, for each channel state j, there is a queue-length threshold. The optimal policy transmits at peak power over channel state j only if the queue length exceeds the threshold, and does not transmit otherwise.

Implementing the optimal policy requires knowing the arrival rate and channel statistics. A possible extension of this work is to study adaptive schemes, which can adjust the parameter settings based on the service rate and current channel state.

The tools developed here could also be applied to study other resource allocation and control problems in wireless networks. For example, one could investigate the optimal scheduler for a multi-class queue and multiple servers with time-varying channels. However. extending the results to multiple queues is hardly straightforward. The main difficulty is that the reflection direction is not fixed but depends on the control policy. This complicates the optimization problem. Concerning existence of an invariant measure and explicit solutions for the density for the multi-dimensional problem see [6, 45].

In the current approach for the time-varying environment, the dynamics of channel variation disappears in the diffusion limit because fast variation assumption is adopted, namely, the time-scale of the channel variation is faster than the time-scale on which the decision is made. It might be interesting to consider the time-varying component lives at the same time scale of decision making, which might lead to a switching diffusion model in the heavy traffic limit.

7.3 Power of side information: cognitive radio channels

In Chapter 6, we have applied an information theoretic approach to investigate the performance limit of cognitive radio channels. This limit is characterized by the capacity region of two-user interference channels with degraded message sets, in which one transmitter knows the other transmitter's message thus it can cooperate with it to boost the capacity region. We have mainly obtained two results on this class of channels: (i) for the general discrete memoryless IFC setting, we have found achievable regions and outer bounds, which meet under additional assumptions; (ii) for the Gaussian IFC setting, we have determined the capacity region of those channels with weak interference.

Possible extensions along this direction include

- cognitive radio channels with strong interference;
- multi-antenna cognitive radio channels, when both transmitters and receivers have multiple transmit antenna;
- interference channels when one transmitter can only partially access the other transmitter's message, i.e., knowing a lossy function of the message $f(W_2)$.
- beyond interference channels, how the knowledge of side information of other transmitters' messages can improve the transmission efficiency.

Appendix A

Proofs for the heavy traffic model and its controlled diffusion limit

A.1 The Heavy-traffic limit

We apply the methodology in [21, Section III], with a slightly different scaling, and obtain the heavy-traffic limit. We consider a sequence of single-queue systems with time-varying channel process $L^n(t) = L(n^{-\kappa}t)$ and define the scaled queue size by

$$x^n(t) = n^{-\frac{1+\kappa}{2}} q(nt)\,.$$

Let

$$\begin{aligned} A^n(t) &:= n^{-\frac{1+\kappa}{2}} \times \text{number of arrival bits by time } nt \\ D^n(t) &:= n^{-\frac{1+\kappa}{2}} \times \text{number of bits transmitted by time } nt\,. \end{aligned}$$

Then the queue dynamics can be described by

$$x^n(t) = x^n(0) + A^n(t) - D^n(t)\,,$$

where the service process $D^n(t)$ is coupled with the power allocation and the channel process. Using (5.3), we obtain

$$\begin{aligned} D^n(t) &= \frac{1}{n^{\frac{1+\kappa}{2}}} \int_0^{nt} \sum_{j=1}^{N} I_{\{L^n(s)=j\}} r(P_n, j) I_{\{x^n(s)>0\}}\, ds \\ &= \frac{1}{n^{\frac{1+\kappa}{2}}} \int_0^{nt} \sum_{j=1}^{N} \left(r_0(j) + \frac{\gamma_j u_j}{n^{\frac{1+\kappa}{2}}} \right) \\ &\qquad \times I_{\{L(n^{-\kappa}s)=j\}} I_{\{x^n(s)>0\}}\, ds \\ &= \frac{1}{n^{\frac{1-\kappa}{2}}} \int_0^{tn^{1-\kappa}} \sum_{j=1}^{N} \left(r_0(j) + \frac{\gamma_j u_j}{n^{\frac{1-\kappa}{2}}} \right) \\ &\qquad \times I_{\{L(s')=j\}} I_{\{x^n(s')>0\}}\, ds'\,. \end{aligned} \tag{A.1}$$

Let

$$M^{d,n}(t) := \frac{1}{n^{\frac{1-\kappa}{2}}} \int_0^{tn^{1-\kappa}} \sum_{j=1}^{N} I_{\{L(s')=j\}} r_0(j)\, ds' - \lambda^a n^{\frac{1-\kappa}{2}} t\,. \tag{A.2}$$

By (5.2), we have

$$M^{d,n}(t) = \frac{1}{n^{\frac{1-\kappa}{2}}} \int_0^{tn^{1-\kappa}} \sum_{j=1}^{N} \Big[I_{\{L(s')=j\}} - \pi_j \Big] r_0(j)\, ds'\,.$$

By Donsker's theorem [32], $M^{d,n}$ converges weakly to a Wiener process w^d with a finite variance σ_d^2, as $n \to \infty$. At the same time, the centered process of arrivals $M^{a,n}(t) := A^n(t) - \lambda^a n^{\frac{1-\kappa}{2}} t$ also converges weakly to a Wiener process w^a with variance σ_a^2. Furthermore, by Assumption 5.3.1, w^d and w^a are independent. Let

$$B^{d,n}(t) := \frac{1}{n^{1-\kappa}} \int_0^{tn^{1-\kappa}} \sum_{j=1}^{N} I_{\{L(s')=j\}} \gamma_j u_j\, ds'\,. \tag{A.3}$$

Then,

$$\begin{aligned} B^{d,n}(t) \xrightarrow[n\to\infty]{} & \int_0^t \sum_{j=1}^{N} \pi_j \gamma_j u_j(s)\, ds \\ &= \int_0^t b\big(u(s)\big)\, ds\,, \quad \text{a.s.,} \end{aligned}$$

by functional law of large numbers (FLLN) [23]. The scaled idle time for the queue with channel state j is

$$T^n(j,t) = \frac{1}{n^{\frac{1+\kappa}{2}}} \int_0^{nt} I_{\{L^n(s)=j\}} I_{\{x^n(s)=0\}}\, ds\,. \tag{A.4}$$

Thus, we define

$$z^n(t) := \sum_{j=1}^{N} r_0(j) T^n(j,t)\,, \tag{A.5}$$

which can be viewed as the scaled number of bits in the queue that could have been transmitted with the power allocation $P_0(j)$. By (A.1)–(A.5),

$$D^n(t) = \lambda^a n^{\frac{1-\kappa}{2}} t + M^{d,n}(t) + B^{d,n}(t) - z^n(t)\,.$$

Thus,

$$\begin{aligned} x^n(t) = x(0) &+ \big(\lambda^a n^{\frac{1-\kappa}{2}} t + M^{a,n}(t)\big) \\ &- \big(\lambda^a n^{\frac{1-\kappa}{2}} t + M^{d,n}(t) + B^{d,n}(t) - z^n(t)\big) \\ = x(0) &- B^{d,n}(t) + M^{a,n}(t) \\ &\qquad - M^{d,n}(t) + z^n(t)\,, \end{aligned} \tag{A.6}$$

Note that $z^n(t)$ is also the *reflection term* of the process $x^n(t)$ (e.g., see [59]), satisfying,

$$\begin{aligned} z^n(t) = \max \Big\{0, -\min_{s\le t} \big[x^n(0) - B^{d,n}(s) \\ + M^{a,n}(s) - M^{d,n}(s)\big]\Big\}\,. \end{aligned} \tag{A.7}$$

By the weak convergence of $M^{a,n}(t)$, $M^{d,n}(t)$ to their continuous limits on the right side of (A.7), $z^n(t)$ thus converges weakly to $z(t)$, where

$$z(t) = \max\left\{0, -\min_{s\le t}\left[x(0) - \int_0^s b\bigl(u(s')\bigr)\,ds' + w^a(s) - w^d(s)\right]\right\}.$$

Thus (A.6) converges weakly to (5.4) by the preceding discussion, where $\sigma^2 = \sigma_a^2 + \sigma_d^2$.

A.2 Proofs of Theorem 5.4.3 and Lemma 5.5.2

We start with some preliminary discussion. Let $\bar{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$ denote the one point compactification of $\mathbb{R}_+$ and let $\bar{\mathcal{G}}$ denote the closure of $\mathcal{G}$ in $\mathcal{P}(\bar{\mathbb{R}}_+ \times \widetilde{U})$. Since $\mathcal{P}(\bar{\mathbb{R}}_+ \times \widetilde{U})$ is compact, so is $\bar{\mathcal{G}}$, and hence any sequence of probability measures $\{\nu_k : k \in \mathbb{N}\}$ in $\mathcal{G}$ contains a subsequence which converges weakly in $\bar{\mathcal{G}}$. Furthermore, using the criterion in (5.9) one can show (see [15]) that any $\nu \in \bar{\mathcal{G}}$ can be decomposed as follows: there exists $\delta \in [0,1]$ and probability measures $\nu' \in \mathcal{G}$ and $\nu'' \in \mathcal{P}(\{\infty\} \times \widetilde{U})$ such that for any Borel set $B \subset \bar{\mathbb{R}}_+ \times \widetilde{U}$,

$$\nu(B) = \delta\nu'\bigl(B \cap (\mathbb{R}_+ \times \widetilde{U})\bigr) + (1-\delta)\nu''\bigl(B \cap (\{\infty\} \times \widetilde{U})\bigr). \tag{A.8}$$

We also make use of the following lemma.

Lemma A.2.1. *Let $\mathfrak{M}_s(\mathbb{R}_+ \times \widetilde{U})$ denote the space of finite signed measures on $\mathbb{R}_+ \times \widetilde{U}$, and let $H_1, \dots, H_n$ be half spaces of the form*

$$H_i = \left\{\nu \in \mathfrak{M}_s(\mathbb{R}_+ \times \widetilde{U}) : \int g_i\,d\nu \le k_i\right\},$$

where $g_i : \mathbb{R}_+ \times \widetilde{U} \to \mathbb{R}_+$ are continuous, and $k_i \in \mathbb{R}_+$, $i = 1, \dots, k$. Suppose $H_i \neq \varnothing$, for $i = 1, \dots, k$, and let $H = H_1 \cap \cdots \cap H_k$. Then $(\mathcal{G} \cap H)_e \subset \mathcal{G}_e$.

The proof of Lemma A.2.1 is contained in [16,31], and relies on the following: It is shown in [31] that the convex set $\mathcal{G}$, when viewed as a subset of $\mathfrak{M}_s(\mathbb{R}_+ \times \widetilde{U})$, does not have any finite dimensional faces other than its extreme points. Since H is the intersection of a finite collection of closed half-spaces in $\mathfrak{M}_s(\mathbb{R}_+ \times \widetilde{U})$, it has finite co-dimension in $\mathfrak{M}_s(\mathbb{R}_+ \times \widetilde{U})$. Hence, there are no extreme points in $\mathcal{G} \cap H$, other than the ones in $\mathcal{G}_e$.

An application of Choquet's Theorem (see [15]), together with Corollary 5.4.2 and Lemma A.2.1 yield the following.

Lemma A.2.2. *Let $\nu \in \mathcal{G} \cap H(\bar{p})$. Then there exists $v \in \mathfrak{U}_{se}$ such that $\nu_v \in H(\bar{p})$ and*

$$\int_{\mathbb{R}_+ \times \widetilde{U}} c(x)\, \nu_v(dx, d\tilde{u}) \le \int_{\mathbb{R}_+ \times \widetilde{U}} c(x)\, \nu(dx, d\tilde{u})\,.$$

We now prove Theorem 5.4.3 and Lemma 5.5.2.

Proof of Theorem 5.4.3: First suppose c is unbounded. Fix $\bar{p} \in (0, p_{\max}]$ and let $\{\nu_k\}$ be a sequence in $H(\bar{p})$ such that

$$\lim_{k\to\infty} \int_{\mathbb{R}_+ \times \widetilde{U}} c\, d\nu_k \to J^*(\bar{p})\,. \tag{A.9}$$

Since c was assumed asymptotically unbounded, it follows that the sequence $\{\nu_k\}$ is tight in $\mathcal{P}(\mathbb{R}_+ \times \widetilde{U})$ and hence converges weakly to some ν^* in $\mathcal{P}(\mathbb{R}_+ \times \widetilde{U})$. Clearly, in view of (A.8), $\nu^* \in \mathcal{G}$. On the other hand, since h is continuous and bounded, and $\nu_k \to \nu^*$, weakly, we obtain

$$\int_{\mathbb{R}_+ \times \widetilde{U}} h\, d\nu^* = \lim_{k\to\infty} \int_{\mathbb{R}_+ \times \widetilde{U}} h\, d\nu_k \le \bar{p}\,.$$

Hence, $\nu^* \in H(\bar{p})$. Since the map $\nu \mapsto \int c\, d\nu$ is lower-semicontinuous on $\mathcal{G}$, we have

$$\begin{aligned} \int_{\mathbb{R}_+ \times \widetilde{U}} c\, d\nu^* &\le \liminf_{k\to\infty} \int_{\mathbb{R}_+ \times \widetilde{U}} c\, d\nu_k \\ &= J^*(\bar{p})\,, \end{aligned}$$

and thus ν^* attains the infimum in (5.15).

Now suppose c is bounded. As before, let $\{\nu_k\}$ be a sequence in $\mathcal{G}$ satisfying (A.9) and let $\tilde{\nu}$ be a limit point of $\{\nu_k\}$ in $\bar{\mathcal{G}}$. Dropping to a subsequence if necessary, we suppose without changing the notation that $\nu_k \to \tilde{\nu}$ in $\bar{\mathcal{G}}$, and we decompose $\tilde{\nu}$ as in (A.8), i.e.,

$$\tilde{\nu} = \delta\, \tilde{\nu}' + (1-\delta)\tilde{\nu}''\,,$$

with $\tilde{\nu}' \in \mathcal{G}$, $\tilde{\nu}'' \in \mathcal{P}(\{\infty\} \times \widetilde{U})$, and $\delta \in [0,1]$. Then, on the one hand

$$\delta \int_{\mathbb{R}_+ \times \widetilde{U}} h \, d\tilde{\nu}' \leq \liminf_{k\to\infty} \int_{\mathbb{R}_+ \times \widetilde{U}} h \, d\nu_k \leq \bar{p}\,, \tag{A.10}$$

while on the other, since c has a continuous extension on $\bar{\mathbb{R}}_+$ (this is a simple consequence of the fact that $\lim_{x\to\infty} c(x)$ exists, and the definition of the topology of the one-point compactification [85]),

$$\begin{aligned} J^*(\bar{p}) &= \lim_{k\to\infty} \int_{\bar{\mathbb{R}}_+ \times \widetilde{U}} c \, d\nu_k \\ &= \delta \int_{\mathbb{R}_+ \times \widetilde{U}} c \, d\tilde{\nu}' + (1-\delta) c_\infty \,. \end{aligned} \tag{A.11}$$

Note that since by Assumption 5.4.1 c is not a constant, $J^*(\bar{p}) < c_\infty$, and hence, by (A.11), $\delta > 0$. Let $\tilde{v} \in \mathfrak{U}_{\mathrm{ss}}$ be the control associated with $\tilde{\nu}'$ and $f_{\tilde{v}}$ be the corresponding density of the invariant probability measure. Let $\hat{x} \in \mathbb{R}_+$ have the value

$$\hat{x} = \frac{1-\delta}{\delta f_{\tilde{v}}(0)}\,,$$

and $v^* \in \mathfrak{U}_{\mathrm{ss}}$ defined by

$$v^*(x) = \begin{cases} 0, & \text{if } x \leq \hat{x} \\ \tilde{v}(x-\hat{x}), & \text{otherwise.} \end{cases}$$

The corresponding density is

$$f_{v^*}(x) = \begin{cases} \delta f_{\tilde{v}}(0), & \text{if } x \leq \hat{x} \\ \delta f_{\tilde{v}}(x-\hat{x}), & \text{otherwise.} \end{cases}$$

By (A.10),

$$\begin{aligned} \int_{\mathbb{R}_+} h\bigl(v^*(x)\bigr) f_{v^*}(x)\, dx &= \delta \int_{\hat{x}}^{\infty} h\bigl(v^*(x)\bigr) f_{\tilde{v}}(x-\hat{x})\, dx \\ &= \delta \int_{\mathbb{R}_+ \times \widetilde{U}} h \, d\tilde{\nu}' \\ &\leq \bar{p}\,. \end{aligned}$$

By construction $f_{v^*}(x) \geq \delta f_{\tilde{v}}(x)$, for all $x \in \mathbb{R}_+$. Hence,

$$\int_{\mathbb{R}_+} c(x)\bigl[f_{v^*}(x) - \delta f_{\tilde{v}}(x)\bigr]\,dx \leq (1-\delta)c_\infty\,. \tag{A.12}$$

By (A.11)–(A.12),

$$\begin{aligned}\int_{\mathbb{R}_+} c(x) f_{v^*}(x)\,dx &\leq \delta \int_{\mathbb{R}_+} c(x) f_{\tilde{v}}(x)\,dx + (1-\delta)c_\infty \\ &= J^*(\bar{p})\,.\end{aligned}$$

Therefore, $v^* \in \mathfrak{U}_{\mathrm{ss}}$ is optimal for (5.15). By Lemma A.2.2, v^* may be selected in $\mathfrak{U}_{\mathrm{se}}$. □

Proof of Lemma 5.5.2: For $\bar{p} \in (0, p_{\max}]$, let $\nu^{(\bar{p})} \in H(\bar{p})$ be an optimal ergodic measure, i.e.,

$$\int_{\mathbb{R}_+\times\widetilde{U}} c\, d\nu^{(\bar{p})} = J^*(\bar{p})\,.$$

Denote by $v^{(\bar{p})} \in \mathfrak{U}_{\mathrm{ss}}$ the associated optimal control, and let $f_{v^{(\bar{p})}}$ stand for the density of the invariant probability measure. Set $\hat{x} = [f_{v^{(\bar{p})}}(0)]^{-1}$, and define $v^* \in \mathfrak{U}_{\mathrm{ss}}$ by

$$v^*(x) := \begin{cases} 0, & \text{if } x \leq \hat{x} \\ v^{(\bar{p})}(x-\hat{x}), & \text{otherwise.} \end{cases}$$

We compute the density of the invariant probability measure as

$$f_{v^*}(x) = \begin{cases} \frac{f_{v^{(\bar{p})}}(0)}{2}, & \text{if } x \leq \hat{x} \\ \frac{f_{v^{(\bar{p})}}(x-\hat{x})}{2}, & \text{otherwise.} \end{cases}$$

Then,

$$\int_{\mathbb{R}_+} h\bigl(v^*(x)\bigr) f_{v^*}(x)\,dx = \frac{\bar{p}}{2}\,.$$

Observe that $f_{v^*}(x) \geq \frac{1}{2} f_{v^{(\bar{p})}}(x)$, for all $x \in \mathbb{R}_+$. Hence, since $c(x) < c_\infty$, for all

$x \in \mathbb{R}_+$, we obtain

$$\begin{aligned} J^*\big(\tfrac{\bar{p}}{2}\big) - \tfrac{1}{2}J^*(\bar{p}) &\leq \int_{\mathbb{R}_+} c(x)\big[f_{v^*}(x) - \tfrac{1}{2}f_{v(\bar{p})}(x)\big]\, dx \\ &< \tfrac{1}{2}c_\infty\,, \end{aligned}$$

which yields the desired result. $\square$

Bibliography

[1] R. Ahlswede, *The capacity region of a channel with two senders and two receivers*, Annals of Probability **2** (1974), no. 5, 805–814.

[2] M. Airy, S. Shakkottai, and Jr. Heath, R. W., *Limiting queueing models for scheduling in multi-user mimo systems*, Proceedings IASTED Conference on Communication, Internet and Information Technology, 2003.

[3] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, *Discrete-time controlled Markov processes with average cost criterion: A survey*, SIAM J. Control Optim. **31** (1993), no. 3, 282–344.

[4] A. Arapostathis, M. K. Ghosh, and V. S. Borkar, *Ergodic control of diffusion processes*, in preparation.

[5] R. Atar, *A diffusion model of scheduling control in queueing systems with many servers*, Ann. Appl. Probab. **15** (2005), no. 1B, 820–852.

[6] R. Atar, A. Budhiraja, and P. Dupuis, *On positive recurrence of constrained diffusion processes*, Ann. Probab. **29** (2001), no. 2, 979–1000.

[7] M. Athans, *On the determination of of optimal costly measurement strategies for linear stochastic systems*, Automatica **18** (1972), 397–412.

[8] D. Avitzour and S. R. Rogers, *Optimal measurement scheduling for prediction and estimation*, IEEE Transactions on Acoustics, Speech, and Signal Processing **38** (1990), no. 10, 1733–1739.

[9] R. F. Bass, *Diffusions and elliptic operators*, Probability and its Applications (New York), Springer-Verlag, New York, 1998.

[10] T. Berger, Z. Zhang, and H. Viswanathan, *The CEO problem*, IEEE Trans. on Info. Theory **42** (1996), 887–902.

[11] R. A. Berry and R. G. Gallager, *Communication over fading channels with delay constraints*, IEEE Transactions on Information Theory **48** (2002), no. 5, 1135–1149.

[12] D. Bertsekas, *Dynamic programming: Deterministic and stochastic models*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1987.

[13] V. I. Bogachev, N. V. Krylov, and M. Röckner, *On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions*, Comm. Partial Differential Equations **26** (2001), no. 11-12, 2037–2080.

[14] V. S. Borkar, *Optimal control of diffusion processes*, Pitman Research Notes in Math. 203, Longman Scientific and Technical, Harlow, UK, 1989.

[15] ———, *Optimal control of diffusion processes*, Pitman Research Notes in Mathematics Series, vol. 203, Longman Scientific & Technical, Harlow, 1989.

[16] ———, *Controlled diffusions with constraints. II*, J. Math. Anal. Appl. **176** (1993), no. 2, 310–321.

[17] ———, *Controlled diffusion processes*, Probability Surveys **2** (2005), 213–244.

[18] V. S. Borkar and M. K. Ghosh, *Controlled diffusions with constraints*, J. Math. Anal. Appl. **152** (1990), no. 1, 88–108.

[19] S. Borst, *User-level performance of channel-aware scheduling algorithms in wireless data networks*, Proceedings IEEE INFOCOM 2003, 2003, pp. 321–331.

[20] M. Bramson, *State space collapse with applications to heavy traffic limits for multi-class queueing networks*, Queueing Systems Theory Appl. **30** (1998), 89–148.

[21] R. Buche and H. J. Kushner, *Control of mobile communications with time-varying channels in heavy traffic*, IEEE Trans. Automat. Control **47** (2002), no. 6, 992–1003, Special issue on systems and control methods for communication networks.

[22] A.B. Carleial, *Outer bounds on the capacity of interference channels*, IEEE Trans. Inform. Theory **29** (1983), 602–606.

[23] H. Chen and D. Yao, *Fundamentals of queueing networks: performance, asymptotics, and optimization*, Applications of Mathematics (New York), vol. 46, Springer-Verlag, New York, 2001, Stochastic Modeling and Applied Probability.

[24] M. H. Costa, *On the Gaussian interference channel*, IEEE Trans. Inform. Theory **31** (1985), 607–615.

[25] M. H. M. Costa, *Writing on dirty paper*, IEEE Trans. Inform. Theory **29** (1983), no. 3, 439–441.

[26] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley Series in Telecommunications, John Wiley & Sons Inc., New York, 1991, A Wiley-Interscience Publication.

[27] I. Csiszar and J. Korner, *Information theory: Coding theorems for discrete memoryless systems*, Akademiai Kiado, Budapest, 1997.

[28] J. G. Dai and J. M. Harrison, *The QNET method for two-moment analysis of closed manufacturing systems*, Ann. Appl. Prob. **3** (1993), 968–1012.

[29] J. G. Dai, Din-Horng Yeh, and Chen Zhou, *The QNET method for re-entrant queueing networks with priority disciplines*, Operations Research **45** (1997), 610–623.

[30] N. Devroye, P. Mitran, and V. Tarokh, *Achievable rates in cognitive radio channels*, IEEE Trans. Inform. Theory **52** (2006), no. 5, 1813–1827.

[31] L. Dubins, *On extreme points of convex sets*, J. Math. Anal. Appl. **5** (1962), 237–244.

[32] R. Durrett, *Probability: theory and examples*, second ed., Duxbury Press, Belmont, 1995.

[33] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov processes*, Grundlehren der mathematischen Wissenschaften, vol. 235, Springer-Verlag, New York, 1979.

[34] N. Elia and S. Mitter, *Stabilization of linear systems with limited information*, IEEE Trans. Automat. Control **46** (2001), no. 9, 1384–1400.

[35] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov models, estimation and control*, Applications of Mathematics, vol. 29, Springer-Verlag, New York, 1995.

[36] S. N. Ethier and T. G. Kurtz, *Markov processes: Characterization and convergence*, John Wiley & Sons, Inc., New York, 1986.

[37] M. I. Freĭdlin, *Diffusion processes with reflection and a directional derivative problem on a manifold with boundary*, Teor. Verojatnost. i Primenen. **8** (1963), 80–88.

[38] A. C. Fu, E. Modiano, and J. N. Tsitsiklis, *Optimal energy allocation and admission control for communications satellites*, IEEE/ACM Transactions on Networking **11** (2003), no. 3, 488–500.

[39] S. S. Ge, Z. Sun, and T. H. Lee, *Reachability and controllability of switched linear discrete-time systems*, IEEE Trans. Automat. Control **46** (2001), no. 9, 1437–1441.

[40] S. I. Gel'fand and M. S. Pinsker, *Coding for channel with random parameters*, Probl. Peredachi Inform. (Probl. Inform. Trans.) **9** (1980), no. 1, 19–31.

[41] I. I. Gihman and A. V. Skorohod, *Stochastic differential equations*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 72, Springer-Verlag, Berlin, 1972.

[42] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, *On a stochastic sensor selection algorithm with applications in sensor scheduling and dynamic sensor coverage*, Automatica **42** (2006), no. 2, 251–260.

[43] T. S. Han and K. Kobayashi, *A new achievable rate region for the interference channel*, IEEE Trans. Inform. Theory **27** (1981), 49–60.

[44] J. M. Harrison and J. A. Van Mieghem, *Dynamic control of brownian networks: State space collapse and equivalent workload formulations*, Ann. Appl. Probab. **7** (1997), 747–771.

[45] J. M. Harrison and R. J. Williams, *Multidimensional reflected Brownian motions having exponential stationary distributions*, Ann. Probab. **15** (1987), no. 1, 115–137.

[46] J. M. Harrison and A. Zeevi, *Dynamic scheduling of a multiclass queue in the Halfin-Whitt heavy traffic regime*, Oper. Res. **52** (2004), no. 2, 243–257.

[47] J.M. Harrison and Vien Nguyen, *The QNET method for two moment analysis of open queueing networks*, Queueing systems **6** (1990), 1–32.

[48] K. Herring and J. Melsa, *Optimum measurements for estimation*, IEEE Trans. on Auto. Control **19** (1974), no. 3, 264–266.

[49] A. Host-Madsen, *A new achievable rate for cooperative diversity based on generalized writing on dirty paper*, Proceedings of IEEE Int. Symp. Inf. Theory, June 2003, p. 317.

[50] S.-P. Hsu, D. M. Chuang, and A. Arapostathis, *On the existence of stationary optimal policies for partially observed MDPs under the long-run average cost criterion*, Systems and Control letters **55** (2006), 165–173.

[51] N. Jindal, U. Mitra, and A. Goldsmith, *Capacity of ad-hoc networks with node cooperation*, Proceedings of IEEE Int. Symp. Inf. Theory, 2004, p. 271.

[52] A. Jovicic and P. Viswanath, *Cognitive radio: An information-theoretic perspective*, Proceedings of IEEE Int. Symp. Inf. Theory, July 2006, pp. 2413–2417.

[53] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*, Information and System Sciences Series, Prentice Hall, Upper Saddle River, New Jersey, 2000.

[54] Ja. A. Kogan, *The optimal control of a non-stopping diffusion process with reflection*, Teor. Verojatnost. i Primenen **14** (1969), 516–522.

[55] J. Korner and K. Marton, *General broadcast channels with degraded message sets*, IEEE Trans. Inform. Theory **23** (1977), no. 1, 60–64.

[56] G. Kramer, *Outer bounds on the capacity of Gaussian interference channels*, IEEE Trans. Inform. Theory **50** (2004), 581–586.

[57] V. Krishnamurthy and R. J. Evans, *Hidden markov model multiarmed bandits: a methodology for beam scheduling in multitarget tracking*, IEEE Trans. on Signal Processing **49** (2001), no. 12, 2893–2908.

[58] N. V. Krylov, *Controlled diffusion processes*, Applications of Mathematics, vol. 14, Springer-Verlag, New York, 1980, Translated from the Russian by A. B. Aries.

[59] H. J. Kushner, *Heavy traffic analysis of controlled queueing and communication networks*, Applications of Mathematics (New York), vol. 47, Springer-Verlag, New York, 2001, Stochastic Modeling and Applied Probability.

[60] H. J. Kushner and D. S. Clark, *Stochastic approximation methods for constrained and unconstrained systems*, Springer Verlag, New York, NY., 1978.

[61] H. J. Kushner and P. G. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Applications of Mathematics, vol. 24, Springer-Verlag, New York, NY, 1992, Stochastic Modeling and Applied Probability.

[62] H. J. Kushner and G. G. Yin, *Stochastic approximation algorithms and applications*, Springer Verlag, New York, NY, 1997.

[63] D. Liberzon, *On stabilization of linear systems with limited information*, IEEE Trans. Automat. Control **48** (2003), no. 2, 304–307.

[64] D. Liberzon and J. P. Hespanha, *Stabilization of nonlinear systems with limited information feedback*, IEEE Trans. Automat. Control **50** (2005), no. 6, 910–915.

[65] D. Liberzon and A. S. Morse, *Basic problems in stability and design of switched systems*, IEEE contr. Syst. Mag. **19** (1999), 59–70.

[66] L. Ljung and T. Soderstrom, *Theory and practice of recursive identification*, M.I.T. Press, Cambridge, MA, 1983.

[67] D. G. Luenberger, *Optimization by vector space methods*, John Wiley & Sons Inc., New York, 1967.

[68] I. Maric, R. D. Yates, and G. Kramer, *The strong interference channel with common information*, Proceedings of 43th Allerton Conference on Communications, Control and Computing, Monticello, IL, Sep. 2005.

[69] ———, *The strong interference channel with unidirectional cooperation*, Proceedings of the UCSD Workshop on Information Theory and its Applications, San Diego, CA, Feb. 2006.

[70] R. Mehra, *Optimization of measurement schedules and sensor designs for linear dynamic systems*, IEEE Trans. on Auto. Control **21** (1976), no. 1, 55–64.

[71] L. Meier, J. Peschon, and R. M. Dressler, *Optimal control of measurement subsystems*, IEEE Trans. Automat. Control **12** (1967), no. 5, 528–536.

[72] B. M. Miller and W. J. Runggaldier, *Optimization of observations: a stochastic control approach*, SIAM Journal of Control and Optimization **35** (1997), no. 5, 1030–1052.

[73] E. Mori and J. J. DiStefano, *Optimal nonuniform sampling interval and test-input for identification of physiological system from very limited data*, IEEE Trans. on Auto. Control **24** (1979), no. 6, 893–900.

[74] G. Nair and R. Evans, *Stabilization with data-rate-limited feedback: Tightest attainable bounds*, Systems Control Lett. **41** (2000), 49–76.

[75] G. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, *Topological feedback entropy and nonlinear stabilization*, IEEE Trans. Automat. Control **49** (2004), no. 9, 1585–1597.

[76] C. Ng and A. Goldsmith, *Transmitter cooperation in ad-hoc wireless networks: Does dirty-paper coding beat relaying?*, Proceedings of IEEE Information Theory Workshop, Oct. 2004.

[77] Y. Oohama, *The rate-distortion function for the quadratic Gaussian CEO problem*, IEEE Trans. on Info. Theory **44** (1998), no. 3, 1057–1070.

[78] ———, *Rate-distortion theory for gaussian multiterminal source coding systems with several side informations at the decoder*, IEEE Trans. on Info. Theory **51** (2005), no. 7, 2577–2593.

[79] E. Plambeck, S. Kumar, and J. M. Harrison, *A multiclass queue in heavy traffic with throughput time constraints: asymptotically optimal dynamic controls*, Queueing Syst. **39** (2001), no. 1, 23–54.

[80] L. K. Platzman, *Optimal infinite-horizon undiscounted control of finite probabilistic systems*, SIAM J. Control Optim. **18** (1980), no. 4, 362–380.

[81] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*, Wiley series in probability and mathematical statistics, John Wiley & Sons, New York, 1994.

[82] T. S. Rappaport, *Wireless communications: Principles and practice*, second ed., Prentice Hall, Englewood Cliffs, 2002.

[83] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, vol. 28, Princeton University Press, Princeton, New Jersey, 1946.

[84] S. M. Ross, *Arbitrary state Markovian decision processes*, Ann. Math. Statist. **39** (1968), no. 6, 2118–2122.

[85] H. L. Royden, *Real analysis*, third ed., Macmillan, New York, 1988.

[86] W. J. Runggaldier and L. Stettner, *Approximations of discrete time partially observed control problems*, Applied mathematics monographs, no. 6, Giardini Editori E Stampatori in Pisa, 1994.

[87] I. Sason, *On achievable rate regions for the Gaussian interference channel*, IEEE Trans. Inform. Theory **50** (2004), 1345–1356.

[88] H. Sato, *Two-user communication channels*, IEEE Trans. Inform. Theory **23** (1977), 295–304.

[89] ______, *The capacity of the Gaussian interference channel under strong interference*, IEEE Trans. Inform. Theory **27** (1981), 786–788.

[90] A. V. Savkin, R. Evans, and E. Skafidas, *The problem of optimal robust sensor scheduling*, Systems Control Lett. **43** (2001), 149–157.

[91] D. Shah, *Randomization and heavy traffic theory: new approaches to the design and analysis of switch algorithms*, Ph.D. thesis, Stanford University, 2004.

[92] M. Shakeri, K.R. Pattipati, and D.L. Kleinman, *Optimal measurement scheduling for state estimation*, IEEE Trans. on Aero. and Elec. Systems **31** (1995), no. 2, 716 – 729.

[93] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla M. Jordan, and S. Sastry, *Kalman filtering with intermittent observations*, IEEE Trans. Automat. Control **49** (2004), no. 9, 1453–1464.

[94] A. L. Stolyar, *Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic*, Ann. Appl. Probab. **14** (2004), no. 1, 1–53.

[95] Z. Sun and S. S. Ge, *Analysis and synthesis of switched linear control systems*, Automatica **41** (2005), no. 2, 181–195.

[96] Z. Sun and D. Zheng, *On reachability and stabilization of switched linear systems*, IEEE Trans. Automat. Control **46** (2001), no. 2, 291–295.

[97] S. Tatikonda and S. Mitter, *Control under communication constraints*, IEEE Trans. Automat. Control **49** (2004), no. 7, 1056–1068.

[98] ______, *Control under noisy channels*, IEEE Trans. Automat. Control **49** (2004), no. 7, 1196–1201.

[99] Joy A. Thomas, *Feedback can at most double Gaussian multiple access channel capacity*, IEEE Trans. Inform. Theory **33** (1987), 711–716.

[100] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, *Energy-efficient packet transmission over a wireless link*, IEEE/ACM Transactions on Networking **10** (2002), no. 4, 487–499.

[101] S. Vishwanath, N. Jindal, and A. Goldsmith, *Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels*, IEEE Trans. Inform. Theory **49** (2003), no. 10, 2658–2668.

[102] P. Viswanath and D. N. C. Tse, *Sum capacity of the vector Gaussin broadcast channel and uplink-downlink duality*, IEEE Trans. Inform. Theory **49** (2003), 1912–1921.

[103] H. Viswanathan and T. Berger, *The quadratic Gaussian CEO problem*, IEEE Trans. on Info. Theory **43** (1997), 1549–1559.

[104] H. Weingarten, Y. Steinberg, and S. Shamai, *The capacity region of the Gaussian MIMO broadcast channel*, IEEE Trans. Inform. Theory **52** (2006), no. 9, 3936–3964.

[105] W. Wong and R. Brockett, *Systems with finite communication bandwidth constraints I: State estimation problems*, IEEE Trans. Automat. Control **42** (1997), no. 9, 1294–1299.

[106] ———, *Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback*, IEEE Trans. Automat. Control **44** (1999), no. 5, 1049–1053.

[107] G. Xie and L. Wang, *Controllability and stabilizability of switched linear-systems*, Systems Control Lett. **48** (2003), no. 2, 135–155.

[108] ———, *Reachability realization and stabilizability of switched linear discrete-time systems*, J. Math. Anal. Appl. **280** (2003), no. 2, 209–220.

[109] G. Xie and L. Wang, *Stabilization of a class switched linear systems*, IEEE International Conference on Systems, Man and Cybernetics, vol. 2, Oct. 2003, pp. 1898–1903.

[110] L. Zhang and D. Hristu-Varsakelis, *LQG control under limited communication*, Proceedings of the 44th IEEE Conference on Decision and Control and the European Control Conference, Dec. 2005, pp. 185–190.

Vita

Wei Wu was born in Harbin, Heilongjiang, China. He received his B.S. in Physics in 1999 and M.S. in Electronic Engineering in 2002, both from Tsinghua University, Beijing, China. Since Fall 2002, he has been pursuing his Ph.D. in Electrical and Computer Engineering at the University of Texas at Austin. His research interests lie in modeling, control and optimization of stochastic systems and analysis of networked systems. In 2006, he received Hemphill-Gilmore Student-endowed Fellowship from UT-Austin.

Permanent Address: 1255 San Tomas Aquino Rd., Apt 113, San Jose, CA 95117

This dissertation was typeset with LaTeX 2_ε[1] by the author.

[1] LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The macros used in formatting this dissertation were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay, James A. Bednar, and Ayman El-Khashab.